

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 24, 2004, 20:12:36 ; Search time 3601 Seconds
(without alignments)
5365.407 Million cell updates/sec

Title: US-10-086-464-2
Perfect score: 3453
Sequence: 1 MSAPSGTGGTSPSPSNT.....REMEMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool_p/US10086464/runat_23042004_083015_5625/app.query.fasta_1.839
-DB=EST -QPMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOQPC=0.1 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10086464 @CGN 1 1 4237 @runat_23042004_083015_5625 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vri:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1092.5	31.6	1084	12	BP184771
2	1090.5	31.6	824	14	CD435142
3	1086	31.5	788	14	CF436437
4	1071	31.0	789	14	CF436655
5	1069	31.0	759	12	BM408099
6	1066	30.9	785	14	CF436122
7	1065	30.8	757	12	BG596561
8	1063	30.8	731	14	CB655196
9	1047	30.3	692	12	BG441204
10	1030	29.8	723	14	CD839231
11	1027.5	29.8	1053	11	AY108241
12	1025	29.7	1016	11	AY108243
13	1018	29.5	666	12	BM358715
14	1013	29.3	695	13	CA072174
15	983	28.5	715	13	BQ099573
16	961	27.8	770	13	BQ999193
17	931	27.0	611	14	CA237156
18	931	27.0	859	29	CC725849
19	929.5	26.9	809	14	CA765135
20	928	26.9	850	29	CC688754
21	923.5	26.7	769	14	CK283399
22	921	26.7	618	14	CA269355
23	921	26.7	651	14	CF478389
24	920	26.6	673	14	CA164704
25	919.5	26.6	938	29	CG333846
26	917.5	26.6	949	29	CG436431
27	912	26.4	679	13	BQ404121
28	904.5	26.2	674	14	CA298046
29	903	26.2	595	14	CF015663
30	903	26.2	637	13	BQ134241
31	903	26.2	645	13	BQ506869
32	897.5	26.0	899	14	CK096498
33	895.5	25.9	682	13	CA095337
34	895	25.9	641	14	CA248316
35	895	25.9	725	29	CG450877
36	893.5	25.9	803	14	CF243440
37	891	25.8	646	13	BQ240617
38	891	25.8	652	13	BQ496883
39	891	25.8	666	13	BQ499754
40	891	25.8	691	14	CA237152
41	889	25.7	626	13	BQ582873
42	888.5	25.7	593	10	BF176907
43	887	25.7	622	14	CA999610
44	887	25.7	673	13	BQ410602
45	886.5	25.7	573	9	AV551753

ALIGNMENTS

RESULT 1
BP184771
LOCUS
DEFINITION BP184771 pns rice panicle cDNA, germ cell generating stage Oryza sativa (japonica cultivar-group) cDNA, mRNA sequence.
ACCESSION BP184771
VERSION BP184771.1
KEYWORDS EST, GI:32948199
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

```

REFERENCE
AUTHORS      Moriguchi, K., Ito, Y., Yamazaki, Y. and Kurata, N.
TITLE        Finding of various plant nuclear proteins using yeast nuclear
              transposition trap system - a proteomal approach
JOURNAL      Unpublished (2003)
COMMENT      Contact: Kazuki Moriguchi
              Plant Genetics
              National Institute of Genetics
              Yata 1111, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-55-981-6872
              Fax: 81-55-981-6879
              Email: kmoriguc@lab.nig.ac.jp
              cDNA clone obtained from nuclear transposition trap system
              encoding a protein similar to Oryza sativa (japonica
              cultivar-group) putative receptor protein kinase PERK1.
FEATURES
source      Location/Qualifiers
            1..1084
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /db_xref="taxon:39947"
            /tissue_type="panicle"
            /dev_stage="germ cell generating stage"
            /clone_lib="PNS rice panicle cDNA, germ cell generating
            stage"
ORIGIN
Alignment Scores:
Pred. No.:      4,98e-26      Length:      1084
Score:          1092.50      Matches:      212
Percent Similarity: 82.88%      Conservative: 30
Best Local Similarity: 72.60%      Mismatches: 43
Query Match:    31.64%      Indels:      7
DB:             12      Gaps:      3
US-10-086-464-2 (1-647) x BP184771 (1-1084)
QY 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
DB 2 TTGCACGGAAGGCGGACCAACATGAGTGCGGCCACACAGCTAAAGATTGCTTTGGGA 61
QY 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
DB 62 GCTGCAAAAGGGTTAGCTTATCTTCATGAAGACTGCCATCTCAAGATCATCCATCGTGAT 121
QY 400 IleLysAlaSerAsnIleLeuLysPheLysPheGluAlaLysValAlaAspPheGly 419
DB 122 ATTAAGGCGTCAAAACATCTCTTGAATTGAATCTAAGGTGCTGATTTTGGGA 181
QY 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
DB 182 CTTGTGAAGTTCACCGATGATAATAACACTCATGTTCGACAAAGAGTAATGGCACTTTT 241
QY 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
DB 242 GGATATCTAGCACCAAGAGTACGATCTTCTGTCGAAGCTCACTGAGAAATCAGATGCTTC 301
QY 460 SerPheGlyValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 479
DB 302 TCCATGAGGATTGCTTCTTGAGTTAATACTGGTGGCCAGTTGATACAAGCA 361
QY 480 ValTyrValAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
DB 362 ACATATATGATGATGACAGCTTGGTTCAGTGGGCAAGGCCTTTACTGATGACAGCACTTGG 421
QY 500 GlnGlyAspPheGluGlyLeuAlaAlaPheAlaLysMetAsnGlyTyrAspArgGluGlu 519
DB 422 AATGGTAACACAGGAGGTAGTAGATCTCGGCTTGGGAGGATTTCAATCCCAATGAG 481
QY 520 MetAlaArgMetValAlaCysAlaAlaAlaAlaCysValArgHisSerAlaArgArgPro 539
DB 482 ATGGGCGAAGTAATGTTGCTGTCAGCTGATGATCGGACCATTCGCTCGTTCGCCCA 541
QY 540 ArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGlu 559

```

```

Db 542 CGCATGAGCCAGGTGTTCGGGCTTTGCAAGGTGACGTGCTTTGGAGGATCTTAATCAA 601
QY 560 GlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAsp 579
DB 602 GGTGTTGGGCTGTGTCACAGCCGTATTTTGGATCGTAC---AGCAGCTCTGACTATGAT 658
QY 580 SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGlu 599
DB 659 TCTGGCCATATACACAGGAGCATGAGAGTTCAGGAAGATGGCTTTT---ACCAACAAT 715
QY 600 TyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGly 619
DB 716 AATGATACGAGCAGTCAATACAGCGCACCAACACGAGTATGCCAGATACCCCTCTGCA 775
QY 620 SerSerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly 639
DB 776 TCAAGCAGCGAGGGCCAAACAAACCCAGGAAGTCAGAGCAGGACCAACCAAGAGAGCGGC 835
QY 640 -----GlnGlyTyrSerGlyProSer 646
DB 836 TACAGTGGCTACAGCTCAGGATACAGCGAGCCTCA 871
RESULT 2
CD435142
LOCUS      CD435142
DEFINITION EL01N0355D03.b EndospERM_3 Zea mays cDNA, mRNA linear EST 03-JUN-2003
ACCESSION  CD435142
VERSION    CD435142.1 GI:31350785
KEYWORDS  EST.
SOURCE     Zea mays
ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 824)
AUTHORS   Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and
            Messing, J.
TITLE     Sequencing of the maize endospERM ESTs
JOURNAL   Unpublished (2002)
COMMENT   Contact: Lai, Jinsheng
            Dr. Joachim Messing's lab
            Wakeman Institute, Rutgers University
            190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
            Tel: 732-445-3801
            Fax: 732-445-5735
            Email: jlai@wakeman.rutgers.edu
            Seq primer: T3.
FEATURES  Location/Qualifiers
            source
            1..824
            /organism="Zea mays"
            /mol_type="mRNA"
            /cultivar="W22"
            /db_xref="taxon:4577"
            /tissue_type="EndospERM of 7-23DAP"
            /clone_lib="EndospERM 3"
            /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
            XhoI"
ORIGIN
Alignment Scores:
Pred. No.:      4.54e-26      Length:      824
Score:          1090.50      Matches:      206
Percent Similarity: 84.12%      Conservative: 27
Best Local Similarity: 74.37%      Mismatches: 41
Query Match:    31.58%      Indels:      3
DB:             14      Gaps:      2
US-10-086-464-2 (1-647) x CD435142 (1-824)
QY 334 LeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsn 353
DB 2 TTGGTTGGCTATTGCAATTTCTGGAGGCACCTTGCTGCTGTGCTATGAGTTTGTCCTCAAT 61

```


Utterback, T., Chiemingo, A., Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.
 Generation of ESTs from potato roots
 Unpublished (2001)
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
 http://genome.arizona.edu/orders/
 Seq primer: T3

Location/Qualifiers
 1. .759
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cPRO33021"
 /tissue_type="roots"
 /dev_stage="in vitro grown stem cuttings"
 /lab_host="SOLR"
 /clone_lib="potato roots"
 /note="vector: pluscript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley lab; sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."
 Alignment Scores:
 Pred. No.: 2,04e-25 Length: 759
 Score: 1069.00 Matches: 200
 Percent Similarity: 89.29% Conservative: 25
 Best Local Similarity: 79.37% Mismatches: 27
 Query Match: 30.96% Indels: 0
 DB: 12 Gaps: 0
 US-10-086-464-2 (1-647) x BM408099 (1-759)

ORIGIN

312 GlyGluArgGluPheGlnAlaGluValGluIleSerArgValHisHisArgHisLeu 331
 Db 2 GGGGAACGTGAATTTTCAGCGGAAGTTGAGATTATTAGCCGAGTACATCAAGCATCTT 61
 332 ValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheVal 351
 Db 62 GTGTCCTTGTGGATCTGCACTTACTTGGGCTCAGAGACTGCTTGTATTGAGTTGTT 121
 352 ProAsnAsnLeuLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTrpSer 371
 Db 122 CCAACCAATACTTTGGAAATTTTCATTACCGGAAGGAGGCCCTCTTGGATTGGCCA 181
 372 ThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys 391
 Db 182 ATACGGCTAAGATTTGCTAGGTGCTAGCTAAAGGACTGCGATATCGCATGAAGACTGC 241
 392 AsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPhe 411
 Db 242 CAACCGAATCATCTCCCGTGATATCAGGCAGCTAATATCTTATCGACTTAATTTT 301
 412 GluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisVal 431
 Db 302 GAGGCTAAGGTTGCTGATTTTGGACTTGCACCAAGCTAATCTCTGATTAATCTCATGTC 361
 432 SerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLys 451
 Db 362 TCCACCAAGAGTGAATGGGAACCTTTGGGTATTGGCTCCAGAAATGCTTCTCTGGAAG 421
 452 LeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuThrGly 471
 Db 422 CTTACACCAAGTCAAGCATCTCTCTCTTTGGTGTAAATGCTTCTTGTGATTAATCTGGA 481

Qy 472 ArgArgProValAspAlaAsnValTyrValAspAspSerLeuValAspTrpAlaArg 491
 Db 482 CGTCGGCCTTGTACTCAATCATCATGAGATAGTTTGGTGGACTGGGCACGT 541
 Qy 492 ProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMet 511
 Db 542 CCAATTACTCACAGAGCTTTAGAGATGAAAAGTTTGTATACCTTGTGTATCGTCGGCTA 601
 Qy 512 AsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysVal 531
 Db 602 GAAATGATTATAACCAATAATGAGTCGCTCGCATGCTTGTCTTGTCTGCTTGTGTG 661
 Qy 532 ArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsn 551
 Db 662 CGTCATTTCAGCAAGCGTAGACCAACGATGACACAGCTTCTCGAGCTTGGAAAGGAT 721
 Qy 552 ValSerLeuSerAspLeuAsnGluGlyMetArgPro 563
 Db 722 GTTCATTATCAGACCTTTTACGAAGGATTAACCT 757

RESULT 6
 CF436122
 LOCUS 785 bp mRNA linear EST 04-SEP-2003
 DEFINITION EST672467 normalized cDNA library of onion Allium cepa cDNA clone
 ACACJ26, mRNA sequence.
 ACCESSION CF436122
 VERSION CF436122.1 GI:34458812
 KEYWORDS EST.
 SOURCE Allium cepa (onion)
 ORGANISM Allium cepa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
 Allium.
 REFERENCE 1 (bases 1 to 785)
 AUTHORS Havey, M.J., Cheung, F., Van Aken, S., Utterback, T. and Town, C.D.
 TITLE Expressed Sequence Tags from a normalized library of mixed onion
 tissues (Allium cepa)
 JOURNAL Unpublished (2003)
 COMMENT Contact: Havey MJ
 Department of Horticulture
 USDA-ARS and University of Wisconsin
 1575 Linden Drive, Madison, WI 53706, USA
 Tel: 608-262-1830
 Fax: 608-262-4743
 Email: mjhavey@facstaff.wisc.edu
 TIGR sequence name ACACJ26TR. For more information:
 http://haveylab.hort.wisc.edu
 Seq primer: CAG GAA ACA GCT ATG ACC.

FEATURES

source
 1. .785
 /organism="Allium cepa"
 /mol_type="mRNA"
 /cultivar="Red Creole (bulbs), unknown(callus), Ebano &
 Texas Legend(roots)"
 /db_xref="taxon:4679"
 /clone="ACACJ26"
 /tissue_type="Callus, roots, and young bulbs"
 /clone_lib="normalized cDNA library of onion"
 /notes="Vector: pCMVSPORT6 1-ccdb (Invitrogen); Site 1:
 EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA
 from callus, roots, and young bulbs were combined to
 synthesize the library. Normalization to enrich for
 low-copy transcripts was performed by proprietary
 techniques of Invitrogen."

ORIGIN

Alignment Scores:
 Pred. No.: 2.61e-25 Length: 785
 Score: 1066.00 Matches: 204
 Percent Similarity: 86.64% Conservative: 23
 Best Local Similarity: 77.86% Mismatches: 34
 Query Match: 30.87% Indels: 1
 DB: 14 Gaps: 0

US-10-086-464-2 (1-647) x CF436122 (1-785)

```
QY 260 LysSerThrPheThrTyGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
Db 2 AAGAGCACTTTACATATAGAGAATTGGCGATAGCTACAAATGGGTTTTCCGACTATAT 61
QY 280 LeuLeuGlyGlnGlyGlyPheGlyValHisGlyValLeuProSerGlyLeuGlu 299
Db 62 CTTCTTGGCAGAGTGGATTGGATATGTGCACAAGGAGTACTTCCAAACGGTHAAGAA 121
QY 300 ValAlaValLeuGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGlu 319
Db 122 GTGGCTATCAAGCAGTGAAGCGGAGCGACAAAGGGAGCGTGAGTTTCAAGCAGAG 181
QY 320 ValGluLeuLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIle 339
Db 182 GTTGAATATATCAGTCGAGTGCATCATAGGCAATTTGGTTTCTTTAGTTGGCTATTGCAAT 241
QY 340 AlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHis 359
Db 242 TCAGGAGATCATAGATTGCTGTATGAATATGTTCTTAATAAAACCCCTTGAGTTCCAT 301
QY 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db 302 TTGCATGGAAGAAAGATGTTCCACCTATGGATTGGCCACCGCGTTAAAAAATGCTTTGGGT 361
QY 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleLeuHisArgAsp 399
Db 362 TCTGCCAAGGGGTTGGCATATCTCCATGAAGATTGTCATCCCAAAATTAATTCATCGTGAT 421
QY 400 IleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGly 419
Db 422 ATTAAGCAGCAAAATATCTCTCGAGATGAACCTTTGAGGCTAAGTTGCGATTTTGGC 481
QY 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 482 CTTGCAAAAGTTTCACTGTTGMAAATAACACCCATGTTTCTACACGAGTCATGGGAACATTT 541
QY 440 GlyTyrLeuAlaProGluTyrAlaLysArgGlyLysLeuThrGluLysSerAspValPhe 459
Db 542 GGGTATCTGGCACCAAGATATGTCATCTCTGGTAAACCTAAGATGAATAATCAGATGCTTT 601
QY 460 SerPheGlyValValLeuLeuLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db 602 TCATTTGGAGTCATGCTTTTGGAGCTGATTACTGGAAAGCGACCACTTGATGCAACCCAA 661
QY 480 ValTyrValAspAspSer-LeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGly 499
Db 662 ACTTATACAGATGATAGTTTGGTCCATTTGGGCAAGACCATTTGCTGCAGAGGCTATGGA 721
QY 499 uGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGly 519
Db 722 GCATGATGATTATGATGAGCTCAITTGACAGAAAGGCTAGGAGATAATAACAACCATGACGA 781
QY 519 uMet 520
Db 782 AATG 785
```

```
RESULT 7
BG596561
LOCUS EST495239 cSTS Solanum tuberosum cDNA clone cSTS15A23 5' sequence,
DEFINITION mRNA sequence.
ACCESSION BG596561
VERSION BG596561.1 GI:13614701
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 757)
```

AUTHORS

van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chierning,A.,
Bougr,O., Buell,C.R., Rinning,C., Tankeley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13F-R.

FEATURES

source

1..757
Location/Qualifiers
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS15A23"
/issue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/clone_lib="cSTS"
/note="Vector: phluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."

ORIGIN

Alignment Scores:
Pred. No.: 2 72e-25 Length: 757
Score: 1065.00 Matches: 200
Percent Similarity: 89.16% Conservative: 22
Best Local Similarity: 80.32% Mismatches: 27
Query Match: 30.84% Indels: 0
DB: 12 Gaps: 0
US-10-086-464-2 (1-647) x BG596561 (1-757)

```
QY 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Db 1 GCAGTTAAACAGCTTAAGCTTGAAGTGCACAAGGGAAACGTGAATTTTCAGCGCGAAGTT 60
QY 321 GluIleLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
Db 61 GAGATTATATAGCCGAGTACATCAACAGCATCTTGTCCTCTTGTGGATCTGCATTACT 120
QY 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu 360
Db 121 GGGGCTCAGAGACTGCTTGTATTATGAGTTGTTCTTCCAAACATACTTTGGAATTTTCATTTA 180
QY 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db 181 CACGAAAGGGAAGGCGCTCTTGGATTGGCAATACGGCTAAAGATTGCTAGAGGTCA 240
QY 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleLeuHisArgAspIle 400
Db 241 GCTAAAGGACTGGCATATCTGCATGAAGACTGCCAACCGAAATCATTCACCGTGATATC 300
QY 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 301 AAGCAGCATATATATCTATTCGACTTTAAATTTTGGGGCTAGGTTGCTGATTTTGACTT 360
QY 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 361 GCCAAGCTAACTTCTGATGTTAATCTCATGCTCTCCACGAGATGATGGAACTTTGGG 420
QY 441 TyrIleAlaProGluTyrAlaLysSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 421 TATTTGGCTCCAGAATATGCTTCTCTGAAAGCTTACAGACAAGTCAGACGATATTCCTC 480
```

```

Qy 461 PheGlyValValLeuLeuGluLeuThrGlyArgArgProValAspAlaAsnVal 480
Db 481 TTTGGTGAATGCTTCTTGATGATTAACATGACGCTGGCTGTGACTCTACTCAATCA 540
Qy 481 TyrValAspSerLeuValAspTIPAlaArgProLeuLeuAsnArgAlaSerGluGln 500
Db 541 TACATCGAAGATAGTTTGGTGGACTGGCAGCTCCATCTACTCACAGGCTTAGAAGAT 600
Qy 501 GlyAspPheGlyLeuAlaAspAlaLeuMetAsnGlyTyrAspArgGluGluMet 520
Db 601 GAAAGATTGATACCTTGTGTGATCGTGGCTAGAAAATGATTATAACCAATAGATG 660
Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 661 GCTCGCATGTTGCTTGTGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 541 MetSerGlnIleValArgAlaLeuGlu 549
Db 721 ATGACACAGGTTCTCCGAGCTTGGA 747

RESULT 8
CB655196 731 bp mRNA linear EST 09-APR-2003
LOCUS OSJNEC08F21.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA
DEFINITION clone OSJNEC08F21 5', mRNA sequence.
ACCESSION CB655196
VERSION CB655196.1 GI:29658921
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
REFERENCE Jantauriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
AUTHORS Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Zhang,G.
TITLE Large-scale identification of ESTs involved in the interaction
JOURNAL between rice and Magnaporthe grisea
COMMENT Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 08 row: F column: 21
Seq primer: gta aaa cga cgg cca gtc.
FEATURES
source
1..731
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEC08F21"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEC"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
ORIGIN
Alignment Scores:
Pred. No.: 3.05e-25 Length: 731
Score: 1063.00 Matches: 199
Percent Similarity: 91.36% Conservative: 23
Best Local Similarity: 81.89% Mismatches: 21
Query Match: 30.78% Indels: 0

```

```

DB: 14 Gaps: 0
US-10-086-464-2 (1-647) x CB655196 (1-731)
Qy 311 GlnGlyGluArgGluPheGlnAlaGluValGluIleSerArgValHisArgHis 330
Db 1 CAGGGAGAGCGTGAATTCAGGGCGGAGTTGAGATTATCAGCAGGTACATCACAAAT 60
Qy 331 LeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuValTyrGluPhe 350
Db 61 CTCGTCTCTCTTGGTTGTTTATTCATTTCTGGGGCAAGAGATTGCTTGTCTATGAGTTT 120
Qy 351 ValProMetAsnLeuLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTyr 370
Db 121 GTCCCCAACACACATTTGGAAATTCACCTTCACGGAAAGCCGACCAATGAGGTGG 180
Qy 371 SerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAsp 390
Db 181 CCCACAGACTAAGATTGCTTTGGGAGCTGCAAGGGTTTAGCTTATCTTCATGAAGAC 240
Qy 391 CysAsnProLysIleIleHisArgAspIleLeuAlaSerAsnIleLeuIleAspPheLys 410
Db 241 TGCCATCCTTAAGATCATCCATGCTGATATTAAAGCGTCAAAACATTTCTTGAATTTAAG 300
Qy 411 PheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHis 430
Db 301 TTTGAATCTAAGGTTGCTGATTTTGGACTTGTAAAGTTCCACAGTGATAATAACACTCAT 360
Qy 431 ValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGly 450
Db 361 GTTTCGACAGAGTAAATGGACATTTTGGATATCTAGCACAGAGTAGCATCTTCTGGC 420
Qy 451 LysLeuThrGlyLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThr 470
Db 421 AAGCTCACTCAGAAATCAGATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Qy 471 GlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuValAspTyrAla 490
Db 481 GGTCTGTCGGCCAGTTGATACAGATCAACATATATGAGTACAGCTTGGTTGACTGGCA 540
Qy 491 ArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAspAlaLys 510
Db 541 AGCCCTTTACTGATGACAGCATCTTGAGATGTTAACTACGAGGAGTTAGTAGATCTCTGG 600
Qy 511 MetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCys 530
Db 601 CTTGGGAAGGATTTCAATCCCATGAGATGCGCAGAGATGATTGCTTGTGAGCTGCATGC 660
Qy 531 ValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGly 550
Db 661 GTACGCCATTTCCGCTCGTCGTCGCCACGATGAGCCAGGTTGTCTCGGGCTTTGGAAGGT 720
Qy 551 AsnValSer 553
Db 721 GACGTGCT 729

RESULT 9
BG441204 692 bp mRNA linear EST 15-MAR-2001
LOCUS BG441204
DEFINITION GA_Ea0012C15f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboeum cDNA clone GA_Ea0012C15f, mRNA sequence.
ACCESSION BG441204.1 GI:13350856
VERSION BG441204.1
KEYWORDS EST.
SOURCE Gossypium arboreum
ORGANISM Gossypium arboreum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 692)
Wing,R.A., Friesch,D., Yu,Y., Main,D., Rambo,T., Simmons,J.,
Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE An integrated analysis of the genetics, development, and evolution

```

of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCATTATAGG
High quality sequence stop: 690.
Location/Qualifiers
1..692
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0012C15f"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 9,38e-25 Length: 692
Score: 1047.00 Matches: 198
Percent Similarity: 94.71% Conservative: 17
Best Local Similarity: 87.22% Mismatches: 12
Query Match: 30.32% Indels: 0
DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BG441204 (1-692)

Qy 247 LeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrglu 266
Db 11 CTTCCGCTGCGTCACCTGGTATTTCCTTAGGTTCTCGAAAGGACCTTTAGCTATGAA 70
Qy 267 GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPhe 286
Db 71 GAATTAGGAGAGCAACGGATGCTCTCGGAAGTTAACTCTCTGGACAAGGTGGTTTT 130
Qy 287 GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys 306
Db 131 GGGTACGTACAAAGGAGTTCTCCCTAATGGGAAGGAGTAGCAGTAAGCACTCAAG 190
Qy 307 ValGlySerGlyGlnGlyArgGluPheGlnAlaGluValGluIleIleSerArgVal 326
Db 191 GCTGGAAGTGGCGAAGCGAGAGAGAAATTCAGGCTGAAGTTGAGATCATTAGCCGCTC 250
Qy 327 HisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeu 346
Db 251 CATCAAAACATCTCGTCTCATTTGGTCGGATACTGTATTTCTGGACAATAAGATGCTT 310
Qy 347 ValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgPro 366
Db 311 GTTTATGAGTTGTTCCAAACAAACACTTGGAGTTTCACTTGCATGGGAAGGGGCGACTG 370
Qy 367 ThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyr 386
Db 371 ACCATGGATTGGCCCAAGGATGAAATTTGCTTTAGGATCTGCAAAAGGAGCTGCATAT 430
Qy 387 LeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 406
Db 431 CTTTCATGAAGATGTTCATCCTAAGATCATTCACCGTGATATTAAAGGCCGCTAATATTCTG 490
Qy 407 IleAspPheLysPheGluAlaLysValAlaAlaAspPheGlyLeuAlaLysIleAlaSerAsp 426
Db 491 TTGGATTTCAAGTTTGAAGCAAGGTGCTGATTTTGGACTAGCCGAAAATGCTTCCGAT 550
Qy 427 ThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyr 446

Db 551 GTCAACACGACGCTCTCCACCAGGCTGATGGGTACTTTTCGGGTATTAGCCCTGAGTAT 610
Qy 447 AlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeu 466
Db 611 GCATTCAAGTGGAAAGCTCACTGATAAATCAGATGTTTCTCTCGGGGTATGCTTTTG 670
Qy 467 GluLeuIleThrGlyArgArg 473
Db 671 GAGTTGATTACCGTCAAGA 691

RESULT 10
CD839231
LOCUS
DEFINITION
CD839231 723 bp mRNA linear EST 10-JUL-2003
RFO2.114104F010529 RFO2 Brassica napus cDNA clone RFO2114104, mRNA
sequence.
ACCESSION
CD839231
VERSION
CD839231.1 GI:32521171
KEYWORDS
EST.
SOURCE
Brassica napus (rape)
ORGANISM
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 723)
Genoplante.
AUTHORS
Genoplante, a major partnership french program in plant genomics
TITLE
Unpublished (2003)
JOURNAL
COMMENT
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.inbioigen.fr>).
Location/Qualifiers
1..723
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="samourai (restored line)"
/db_xref="taxon:3708"
/clone="RFO2114104"
/tissue_type="anthers"
/clone_lib="RFO2"

ORIGIN
Alignment Scores:
Pred. No.: 3,37e-24 Length: 723
Score: 1030.00 Matches: 199
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.83% Indels: 0
DB: 14 Gaps: 0

US-10-086-464-2 (1-647) x CD839231 (1-723)

Qy 449 SerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuLeu 468
Db 2 AGCGAAAGCTCAGGAGAGTCTGACGCTTTCTCATTTGGGTGCTTTTGAGCTC 61
Qy 469 IleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspSerLeuValAsp 488
Db 62 ATTACCGGCGCTCGACCCGTTGATGTCACCAATGCTATGATGACAGCTTAGTTGAC 121
Qy 489 TrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAsp 508
Db 122 TGGGCACACCATTTGCTTAACCGAGCATCTGAGCAAGGAGACTTTGAGGGTTTACCTGAT 181
Qy 509 AlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaIa 528
Db 182 GCAAGATGAATAATGGGTATGACAGAGAGATGGCTCGCATGTTGTTGCTGCTCGG 241

Qy 529 AlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeu 548
 Db 242 GCTTGTGTGCGCATTCAGCTCGCGCAGACCTCGCATGAGCCAGATTGTGCGTGGTTA 301
 Qy 549 GluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnVal 568
 Db 302 GAAGGAATGTATCATCTGTCTGATCTTTAAACGAGGATGAGACCGATCAAGCAATGTA 361
 Qy 569 TyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLys 588
 Db 362 TACAGCTCATACGAGGAGCACCAGATTATGACTCGAGCCAGTACATGAATGAACATGA 421
 Qy 589 LysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsn 608
 Db 422 AAGTTTAGGAAATGGCTCTTTGGAACTCAAGAGTACAAACCCAGCGGTGAGTACAGTAAT 481
 Qy 609 ProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnGlyGlnThrThrArg 628
 Db 482 CCGACGACGACTATGACTGTACCTCTGGTTCAAGCAGCGAGGCGCAACACACAGCGC 541
 Qy 629 GluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
 Db 542 GAAATGAGATGGGGAAGATTAAAGAGAACCGGTGAGGTTATAGTGGACCTTCTCTT 598

RESULT 11
 AY108241 1053 bp mRNA linear HTC 16-OCT-2002
 LOCUS
 DEFINITION Zea mays PC0134818 mRNA sequence.
 ACCESSION AY108241
 VERSION AY108241.1 GI:21211319
 KEYWORDS HTC.

SOURCE
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD Clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 1053)
 Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
 Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
 Unpublished (2002)
 2 (bases 1 to 1053)
 Coe, E.H.
 Direct Submission
 Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

COMMENT
 If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES
 source
 1..1053
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:637889"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Consensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Alignment Scores: 5.61e-24 Length: 1053
 Pred. No.: 1027.50 Matches: 203

Percent Similarity: 79.93% Conservative: 28
 Best Local Similarity: 70.24% Mismatches: 46
 Query Match: 29.76% Indels: 12
 DB: 11 Gaps: 4
 US-10-086-464-2 (1-647) x AY108241 (1-1053)
 Qy 365 ArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLe 384
 Db 6 CGTCCG-----GAGTGGCCTGCTAGATTAAAGATCATGTTGGGTCTGCCAAGGGTTT 59
 Qy 384 uSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAs 404
 Db 60 AGCTTATCTTCATGAGACTGCCATCCAAAGATCATCCATCGTGACATAAAGGCATCTCAA 119
 Qy 404 nIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAl 424
 Db 120 CATCTCTTGGACTTCCAAATTGANGCTAAGTTGCTGATTTTGGACTTGCAGATTCAC 179
 Qy 424 aSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaPr 444
 Db 180 TACTGATAACAACACCCATCTTTTCGACAGAGTAATGGCACCTTTGGGTATTTGGCACC 239
 Qy 444 oGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValVa 464
 Db 240 TGAGTATGCATCTTCTGGCAAGCTAAACAGAAAATCCGATGTATTTTCTTCGGAGTCAT 299
 Qy 464 lLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAs 484
 Db 300 GCTTCTTGACTTATTACTGGGGCGGACCAAGTTGACACACCCCAACATATATGGATGA 359
 Qy 484 pSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGl 504
 Db 360 CAGCTTGGTTGACTGGGCAAGGCCATTTACTGATGCGAGCACTTGAGGATGGTGAATAGA 419
 Qy 504 uGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMetAlaArgMetVa 524
 Db 420 TGCTTTAGTCGATCCTCGGCTGGGAAAGGACTTCAATCTTAATGAGATGCAAGAATGAT 479
 Qy 524 lAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnI 544
 Db 480 AGCTGTGACAGTGCATGTGACGCCATCTGCACGCTCGTCGGCCACGATGATGATCAGGT 539
 Qy 544 eValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGl 564
 Db 540 CGTTCGGGCTTTGGAGGCAATGTCTTTGGAGGACCTTAATGAAGGTGTTTCGGCCTGG 599
 Qy 564 yGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAs 584
 Db 600 CCATAGCCGCTTCTTTGGGTCTATAC---AGCAGCTCCGATTACGATTTCTGGCCAGTCAA 656
 Qy 584 nGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGl 604
 Db 657 CGAGGACATGAAGAGTTCAAGAGATGGCATTCACACACAC-----TATACCAGCAG 710
 Qy 604 yGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGluGl 624
 Db 711 CCAATACAGCGCGCAACACCAAGTGAATATGACAGATACCGCTCTGCATCAAGCAGCGAGG 770
 Qy 624 yGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly----- 639
 Db 771 CCACCAAGACCAAGAGATGGAGTGGGTGCAATGAAGAAAGGTGGTACAGTGGTGCTA 830
 Qy 640 ----GlnGlyTyrSerGlyProSer 646
 Db 831 CAGCTCAGGATACAGCGACCTCG 855
 RESULT 12
 AY108243 1016 bp mRNA linear HTC 16-OCT-2002
 LOCUS
 DEFINITION Zea mays PC0134814 mRNA sequence.
 ACCESSION AY108243
 VERSION AY108243.1 GI:21211321

KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.
AUTHORS 1 (bases 1 to 1016)
 Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 1016)
AUTHORS Coe, E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.
FEATURES
 source
 1..1016
 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Consensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"
ORIGIN
 Alignment Scores:
 Pred. No.: 6,53e-24 Length: 1016
 Score: 1025.00 Matches: 200
 Percent Similarity: 82.14% Conservative: 30
 Best Local Similarity: 71.43% Mismatches: 42
 Query Match: 29.68% Indels: 8
 Db: 11 Gaps: 4
 US-10-086-464-2 (1-647) x AY108243 (1-1016)
 QY 373 ArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsn 392
 Db 3 AGATTAAAGATCGCTCTGGTGCTGCCAAGGTTTAGCTTATCTTCATGAGACTGCCAT 62
 QY 393 ProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu 412
 Db 63 CCARAGATCATCCCGCGACATTAAGGATCTAAGCATCTTCTTGACTTCAAAATTTGAA 122
 QY 413 AlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSer 432
 Db 123 GCTATGGTTGCTGACTTTGGACTTGCAGATTCACCTACTACTATAACACACCATGTGCA 182
 QY 433 ThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 452
 Db 183 ACAAGAGTATGGGACCTTTGGGTATTGGCACCGAGTATGCAGCATCTGGCAAGCTC 242
 QY 453 ThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 472
 Db 243 ACAGAAAAATCCGATGATTCTTTCTTCGAGTCATGCTCTTTGAGCTTATTACTGGCGG 302
 QY 473 ArgProValAspAlaAsnValTyrValAspAspSerLeuValAspTrpAlaArgPro 492
 Db 303 CGACCAATTGACACACCAACCAACATATATATGGATGACAGCTTGGTTGACTGGCGAAGCCA 362

QY 493 LeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsn 512
 Db 363 TTACTGATGAGAGCGCTCGAGATGGTGAATATGATGCTTTGGTGGATCCTCGGTCGGA 422
 QY 513 AsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArg 532
 Db 423 AAGACTTCAATCTTAACGAGATGCGAAGATGATAGCTGTGGGGCTCGCTGGCTAGCG 482
 QY 533 HisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552
 Db 483 CATTCGCGACGCTGTCGGCCACGAATGAGTCAGGTCTGGGCTCTCGAAGGCGACGTG 542
 QY 553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyr 572
 Db 543 TCTTTGGAGGACCTTAATGAAGGTGTCGACCTGCCATAGCCGCTTCTTTGGGTCAAT 602
 QY 573 GlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLys 592
 Db 603 ---AGCAGCTCTGATTACGATTCGGGCGAGTACAAACGAGGACATGCAGAGTTCAGGA 659
 QY 593 MetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAsp 612
 Db 660 ACGGCTTCAATAACACACCTACACGACGAGC---CAATACAGCGCCCAACAGTGA 716
 QY 613 TyrGlyLeuTyrProSerGlySerSerSerGlyGlnThrThrArgGluMet---Glu 631
 Db 717 TACGGCCAGGTGCGCTCTGGATCAAGCAGCGAGGCGCCGACGAGAGATGGAGGAG 776
 QY 632 MetGlyLysIleLysArgThrGly-----GlnGlyTyrSerGlyProSer 646
 Db 777 TCGGTGTCAGTGAAGAAAGGTGCTACACGCGCTACAGCTCCGGATACAGCGGAGCCTCG 836
 RESULT 13
 BM358715 666 bp mRNA linear EST 09-JAN-2002
 LOCUS GA_Ea0012D16r Gossypium arboreum 7-10 dpa fiber library Gossypium
 DEFINITION arboreum cDNA clone GA_Ea0012D16r, mRNA sequence.
 ACCESSION BM358715
 VERSION BM358715.1 GI:18099461
 KEYWORDS EST.
 SOURCE Gossypium arboreum
 ORGANISM Gossypium arboreum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 REFERENCE 1 (bases 1 to 666)
 AUTHORS Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.
 TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber
 JOURNAL Unpublished (2000)
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total High Quality bases = 646
 Seq primer: TAATACGACTCATATAGGG
 High quality sequence stop: 666.
 Location/Qualifiers
 1..666
 /organism="Gossypium arboreum"
 /mol_type="mRNA"
 /strain="AKA"
 /cultivar="8400"
 /db_xref="taxon:29729"
 /clone="GA_Ea0012D16r"
 /tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
 /lab_host="E. coli"

/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Alignment Scores:
Pred. No.: 7,55e-24 Length: 666
Score: 1018.00 Matches: 191
Percent Similarity: 95.41% Conservative: 17
Best Local Similarity: 87.61% Mismatches: 10
Query Match: 29.48% Indels: 0
DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BM358715 (1-666)

Qy 247 LeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGlu 266
Db 11 CTTCCGCTCGTCACCTGGTATTCTCTAGTTTCTCGAAAGCACTTTTGGCTATGAA 70

Qy 267 GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPhe 286
Db 71 GAATTAGCGAGACGAGTGGCTTCTCGAAGTTAACTTCTTGGACAGGTGGTTT 130

Qy 287 GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys 306
Db 131 GGGTACGTACCAAGAGTTCTCCCTATGGGAAGAGTAGCAGTAAGCAACTCAAG 190

Qy 307 ValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgVal 326
Db 191 GCTGGAAGTGGCAAGCGAGAGAGATTTTCAGCTGAAGTTGAGATCATTTAGCCGGCTC 250

Qy 327 HisHisArgHisLeuValSerLeuValGlyTyrCysLeuAlaGlyAlaLysArgLeuLeu 346
Db 251 CATCACAAACATCTCGTCTCATTTGGTCGATCTGTTTCTGGGACAAATAAGATGCTT 310

Qy 347 ValTyrGluPheValProAsnAsnAsnLeuLeuHisLeuHisGlyGluGlyArgPro 366
Db 311 GTTTATGAGTTTGTTCACACACACCTTGGAGTTTCATTGCAATGGAGGGCGACTG 370

Qy 367 ThrMetGluTrpSerThrArgLeuLysLeuAlaLeuGlySerAlaLysGlyLeuSerTyr 386
Db 371 ACCATGATTGGCCGACAGGATGAAATTCCTTTAGGATCTGCAAAAGGACTGGCATAT 430

Qy 387 LeuHisGluAspCysAsnProLysLeuLeuHisArgAspLeuLysAlaSerAsnLeu 406
Db 431 CTTTCATGAAGATTCTCATCTTAAGATCATTCACCGTGATATTAAGGCGCTAATATCTG 490

Qy 407 IleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysLeuAlaSerAsp 426
Db 491 TTGGATTCAAGTTTGAAGCAAGGTTGCTGATTTGGACTAGCGAAATTTGCTTCCGAT 550

Qy 427 ThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyr 446
Db 551 GTCAACAGCGACGTTCTCCACGAGGTGATGGTACTTTCCGGTATTTAGCCCTGAGTAT 610

Qy 447 AlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValVal 464
Db 611 GCITTCAGTGGAAAGCTCACTGATAAATCAGATGTTTCTCTTCGGGGTCATG 664

RESULT 14
CA072174
LOCUS CA072174 695 bp mRNA linear EST 23-SEP-2003
DEFINITION SCCAM1004A11.g AM1 Saccharum officinarum cDNA clone SCCAM1004A11
5', mRNA sequence.
ACCESSION CA072174
VERSION CA072174.1 GI:34924325
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
AUTHORS

TITLE The libraries that made SUBEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 004 row: A column: 11
Seq primer: T7 Promoter Primer.
Location/Qualifiers
source 1..695
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCAM1004A11"
/lab_host="DH10B"
/clone_lib="AM1"
/note="Organ: Apical meristem and tissues surrounding of
mature plants; Vector: pSport1; Site_1: SalI; Site_2:
NotI; An unidirectional cDNA library generated from
[Apical meristem and tissues surrounding of mature
plants]. cDNA was prepared from polyA+ mRNA using
SuperScript plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucet.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:
Pred. No.: 1.13e-23 Length: 695
Score: 1013.00 Matches: 195
Percent Similarity: 91.85% Conservative: 19
Best Local Similarity: 83.69% Mismatches: 16
Query Match: 29.34% Indels: 3
DB: 13 Gaps: 1

US-10-086-464-2 (1-647) x CA072174 (1-695)

Qy 271 AlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHis 290
Db 4 GCACCAATGGTCTTCTGACGCTATCTCTGGGCAAGCGGTTTGGTTGTTTCAC 63

Qy 291 LysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGly 310
Db 64 AAGGAGTGTACCAATGGCACAGAGTTGCTTAAAGCAGTTAAGAGATGGAAGTGGC 123

Qy 311 GlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisHisArgHis 330
Db 124 CAGGGAAAGCCGAGTTTCAGGCAGAGGTTGAGATTATCAGCAGAGTACATCACAGCAT 183

Qy 331 LeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPhe 350
Db 184 CTTGTATCTTGGTGGCTATTGCAATTTCTGGAGCCAAAGTTGCTCTGCTACAGTTT 243

Qy 351 ValProAsnAsnLeuLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTyr 370
Db 244 GTTCCAAACAATACATTGGAATTCACCTTACATGGAAGAGCAGACCAACCTTGGATTGG 303

Qy 371 SerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAsp 390
Db 304 CCAGCAAGACTAAAAATTCCTGGGTTCGCGAAGGATTTGGCATATCTTCATGAAGAT 363

Qy 391 CysAsnProLysIleLeuHisArgAspIleLysAlaSerAsnLeuLeuLeuAspPheLys 410
Db 364 TGGCATCTCAAGATCATTCATCTGTGACATTAAGGCCTCAATATCTTCTTGATCTAAGA 423

Qy 411 PheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHis 430
 Db 424 TTTGAAGCTAAGTGGCAGATTTTGGACTTGCAAAATTCACCTTCTGATACAAACACCCAT 483
 Qy 431 ValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGly 450
 Db 484 GTTTCACACAGAGTAATGGCAGATTTGGTACCTAGCACCTGAGTATGCTGCTTCTGGC 543
 Qy 451 LysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThr 470
 Db 544 AAGCTCACTGAGAAATCAGATGCTCTCTTTGGAGTAATGCTCTTGGAGCTAATAACT 603
 Qy 471 GlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAla 490
 Db 604 GGGCGGCTCT- - - - -GGTAATTCAGACAGGCGAGATGACAACCTTGGGTGACTGGGCA 657
 Qy 491 Arg-ProLeuLeuAsnArgAlaSerGluGlnGlyAsp 502
 Db 658 AGGGCCTTTGATGAATAAAGCATTTGAGGATGGTAAT 694

RESULT 15
 BU099573 715 bp mRNA linear EST 29-AUG-2002
 LOCUS WHE3309_A09_A17S Chinese Spring wheat drought stressed root cDNA
 DEFINITION library Triticum aestivum cDNA clone WHE3309_A09_A17, mRNA
 sequence.
 ACCESSION BU099573
 VERSION BU099573.1 GI:22547372
 KEYWORDS Triticum aestivum (bread wheat)
 SOURCE Triticum aestivum
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 715)
 Anderson, O.D., Chao, S., Close, T.J., Crossman, C., Penton, R.D.,
 Lazo, G.R., Nguyen H.T., Pham, J., Rausch, C.J., Turuspekov, Y.,
 Wilson, C., Woo, J. and Zhang, D.
 The structure and function of the expressed portion of the wheat
 genomes - Chinese Spring drought stressed root cDNA library
 Unpublished (2002)
 Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105595773
 Fax: 5105595818
 Email: oanderson@pw.usda.gov
 Sequences have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: SK primer.

FEATURES
 source
 Location/Qualifiers
 1..715
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="WHE3309_A09_A17"
 /tissue_type="root"
 /dev_stage="Full tillering stage"
 /lab_host="E. coli SOLR"
 /clone_libs="Chinese Spring wheat drought stressed root
 cDNA library"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
 Site_1: EcoRI; Site_2: XhoI; Plants were grown under
 normal conditions, then drought stressed to 80%, 70% and
 60% RWC at Texas Tech University (D. Zhang in HT Nguyen
 lab). Total RNA was prepared separately for roots
 collected at the three different drought conditions. Equal
 amounts of total RNA were pooled from all three samples,
 poly(A) RNA were purified, one cDNA library was made, and
 the cDNA clones were in vivo excised to give pBluescript
 SK(-) phagemids in the TJ Clouse lab at the University of

ORIGIN

Alignment Scores:
 Pred. No.: 1.04e-22 Length: 715
 Score: 983.00 Matches: 184
 Percent Similarity: 89.70% Conservative: 25
 Best Local Similarity: 78.97% Mismatches: 24
 Query Match: 28.47% Indels: 0
 DB: 13 Gaps: 0

US-10-086-464-2 (1-647) x BU099573 (1-715)

Qy 276 SerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuPro 295
 Db 11 TCCGATGCTAATCTCTCGGCAAGGTGGCTTTTGGATTTGTTCAAAGAGGTCTGCCA 70
 Qy 296 SerGlyLysGluValAlaValLysGlnLysValGlySerGlyGlnGlyGluArgGlu 315
 Db 71 GATGGCACAGAGTTGCTGTGAAGCAATTAGAGATGGAAGTGGCAGGAGAGCGTGAG 130
 Qy 316 PheGlnAlaGluValGluIleIleSerArgValHisArgHisLeuValSerLeuVal 335
 Db 131 TTCAGGCGAGAGGTTGAGATTATCAGCGAGTACATCAATAACATCTCGTGACATGGTT 190
 Qy 336 GlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsn 355
 Db 191 GGTATTGCAATTTCTGAAGACAGAGGTTGCTGTCTATGAGTTGTTCCTCCCAATACACG 250
 Qy 356 LeuGluLeuHisLysGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLys 375
 Db 251 TTAGAATTCATATACATGGAGCGGTGGACCACTATGACTGGCCTTCAAGACTACGT 310
 Qy 376 IleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIle 395
 Db 311 ATTGCTTTGGTTCTCGCAAGGATTTGGCGTATCTTCACGAAGACTGCCATCCAAAGATC 370
 Qy 396 IleHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLysVal 415
 Db 371 ATTCATCGTGACATAAAGGCCATCAAAATATTCTCTGGATTACAGATGTGAAGCTAAGGTG 430
 Qy 416 AlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgVal 435
 Db 431 GCAGATTTGGACTTGCAGAGTTAACCTCTGATTAATACACTCATGTTTCCACCAGAGTA 490
 Qy 436 MetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLys 455
 Db 491 ATGGGCACATTTGGGTACCTTGCCACAGAGTATGCTTCTTGGCAGAGCTAACTGAGAAA 550
 Qy 456 SerAspValPheSerPheGlyValValLeuLeuIleThrGlyArgArgProVal 475
 Db 551 TCAGATGCTTTTCTTTTGGAGTAATGCTTCTCGAGTTAATAACTGGCGCGCTCTCTGTA 610
 Qy 476 AspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsn 495
 Db 611 AGTTCAAAACACAGCGCATATGATGACAGCTTGGTTGACTGGCAAGGCCCTTTGATGACA 670
 Qy 496 ArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAsp 508
 Db 671 CAAGCACTCGAGGATGGTAATCAACAGCATGCTTTTAGTGGAT 709

Search completed: April 25, 2004, 00:59:49
 Job time : 3628 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 24, 2004, 20:09:36 ; Search time 5876 Seconds
(without alignments)
4772.455 Million cell updates/sec

Title: US-10-086-464-2
Perfect score: 3453
Sequence: 1 MSSAPSGTGGSPSPSPNST.....REMEMGKIKRTGGGSGPSL 647

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgm2 1/USPTO.spool.p/US10086464/runat 23042004.083015.5613/app query.fasta_1.839
-DB=GenEmbl -QPMF=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10086464 @CGN 1 5265 @runat 23042004.083015.5613 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

RESULT 1

ALIGNMENTS

Result No.	Score	Query Match %	Length	DB ID	Description
1	3453	100.0	1944	6	AX088876 Sequence
2	3453	100.0	1944	6	AX825703 Sequence
3	3453	100.0	2189	6	AX825705 Sequence
4	3453	100.0	2189	8	AY028699 Brassica
5	2936.5	85.0	1959	6	AX825738 Sequence
6	2936.5	85.0	1959	8	BT008400 Arabidops
7	2936.5	85.0	2098	8	AY128792 Arabidops
8	2936.5	85.0	2116	8	AY056788 Arabidops
9	2936.5	85.0	2188	8	AY059901 Arabidops
10	2936.5	85.0	2190	8	AY093065 Arabidops
11	2934.5	85.0	1959	8	BT008409 Arabidops
12	2934.5	85.0	2257	8	AF370509 Arabidops
13	2885	83.6	2324	8	AY089024 Arabidops
14	2622	75.9	79706	8	AB020746 Arabidops
15	1841.5	53.3	2307	8	AK103247 Oryza sat
16	1821.5	52.8	1812	6	AX825735 Sequence
17	1812	52.5	80393	8	AF000382 Arabidops
18	1748.5	50.6	1902	6	AX088882 Sequence
19	1748.5	50.6	1902	6	AX825712 Sequence
20	1748.5	50.6	1939	6	AX088881 Sequence
21	1748.5	50.6	1939	6	AX825711 Sequence
22	1746	50.6	2013	8	BT005955 Arabidops
23	1746	50.6	2320	8	AK118488 Arabidops
24	1687	48.9	2025	6	AX088885 Sequence
25	1687	48.9	2025	6	AX825715 Sequence
26	1687	48.9	2104	6	AX088884 Sequence
27	1687	48.9	2104	6	AX825714 Sequence
28	1668	48.3	1515	6	AX825736 Sequence
29	1637.5	47.4	1488	6	AX825737 Sequence
30	1574.5	45.6	1185	8	AY224526 Oryza sat
31	1567.5	45.4	2364	8	AK064397 Oryza sat
32	1567.5	45.4	2371	8	AK102832 Oryza sat
33	1507	43.6	82356	8	AC006135 Arabidops
34	1507	43.6	92861	8	AC006439 Arabidops
35	1480.5	42.9	125021	8	AC007504 Arabidops
36	1479.5	42.8	114367	8	ATFL0M10 Arabidops
37	1479.5	42.8	125502	8	ATFL0M20 Arabidops
38	1479.5	42.8	195921	8	ATCHRIV81 Arabidops
39	1459.5	42.3	2158	8	AY113877 Arabidops
40	1459.5	42.3	3064	8	AY035076 Arabidops
41	1456	42.2	2046	8	AY113039 Arabidops
42	1456	42.2	2616	8	AF424623 Arabidops
43	1456	42.2	2735	8	AY075681 Arabidops
44	1434.5	41.5	120465	2	AK137001 Oryza sat
45	1431	41.4	3070	8	AK110196 Oryza sat

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Db 841 TTAGCACAGCGGGTTCGTTACGTGTCACAAAGAGTGTGTTGCTAGTGGGAAAGAGTT 900
Qy 301 AlavalysGluLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Db 901 GCTGTGAAGACAGTTGAAGTTGGAGTGGTTCAGGAGAGAGAGGAGTTTCAGGCAGAGGTT 960
Qy 321 GluLeuLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
Db 961 GAGATCATCAGCAGAGTTCACACAGGCATCTGGTGTCTCTTGTGGTTATTGCATCGCC 1020
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu 360
Db 1021 GGTGCCAAAGAGTTGCTGTCTATGAGTTGTGTTCTTAAACAATCTCGAGCTTCACCTC 1080
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db 1081 CATGGCGAGGAGCGCTTACATGAATGGACACAGATTCGAATGCTCTTGGATCT 1140
Qy 381 AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 1141 GCTAAGGACCTTCTTATCTCATGNAGATTGCATCTCTAAATCATTCACCGTGATTC 1200
Qy 401 LysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 1201 AAGGCTTCAACATATTCATAGATTCAAGTTTGAAGCTTAAGGTTGCTGATTTTGGTCTT 1260
Qy 421 AlaLysIleAlaSerAspThrSerHisValSerThrArgValMetGlyThrPheGly 440
Db 1261 GCTAAGATTGCTTCTGATCAAAACACGATGATCAACACGTTGATGGAAACCTTTGGG 1320
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 1321 TACTTGGCTCCGGATACGCTCAAGCGAAGCTCACGGAGAAGTCTGACGTTTCTCA 1380
Qy 461 PheGlyValValLeuLeuLeuLeuThrGlyArgProValAspAlaAsnVal 480
Db 1381 TTTGGCGTTGCTTTTGGAGCTCATTAAGTACGTCGACCGTTCGACCGGTTGATGCCAACATGTC 1440
Qy 481 TyrValAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500
Db 1441 TATGTAGATGACAGCTTAGTTGACTGGGCACCACTTCTTAACCGAGACTCTGAGCAA 1500
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520
Db 1501 GGAGACTTTGAGGTTTACCTGATCAAGATGAATAATGGGTATGACAGAGAGAGATG 1560
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1561 GCTCGCATGGTTGCTGTGCTGGGCTTGTGTTCCGCAATTCAGCTCGCCGACAGCTCGC 1620
Qy 541 MetSerGlnIleValArgAlaLeuGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1621 ATGAGCCAGATTGTGGCTGGTTAGAGGAATGATCATCTGATGATCTTAACGAAGG 1680
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 1681 ATGAGACCAAGTCAAGCAATGTATACAGCTCATACGGAGGAAGCACCGATTATGACTCG 1740
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1741 AGCCAGTACAAATGAAGACATGAAGAAGTTTAGGAAATGGCACTTGGAACTCAAGAGTAC 1800
Qy 601 AsnAlaThrGlyLysSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1801 AACGCCAGGGTGAGTACAGTAAATCCGACCATGACTATGGACTGTACCCGCTGTGGTTCA 1860
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1861 AGCAGCGAGGGCCAAACACACCGAAATGGAGATGGGGAAGATTAAAGAGAACCGGTACG 1920
Qy 641 GlyTyrSerGlyProSerLeu 647
Db 1921 GGTATAGTGGACCTTCTCTT 1941

RESULT 3
LOCUS AX825705 2189 bp DNA linear PAT 11-DEC-2003
DEFINITION Sequence 3 from Patent WO03072763.
ACCESSION AX825705
VERSION AX825705.1 GI:39751232
KEYWORDS Brassica napus (rape)
SOURCE ORGANISM
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 Goring, D., Silva, N. and Haffani, Y. Z.
Increasing plant seed production
Patent: WO 03072763-A 3 04-SEP-2003;
Goring, Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, Z. (CA)
Location/Qualifiers
1..2189
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"
ORIGIN
Alignment Scores:
Pred. No.: 4,38e-76 Length: 2189
Score: 3453.00 Matches: 647
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-086-464-2 (1-647) x AX825705 (1-2189)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerSerSerSerThr 20
Db 97 ATGTCTCGCGCGCTCTCCGGGAGTGGTTCGCTCCATCTCCACCATCAAACTCCACA 156
Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProProPro 40
Db 157 ACCACCATCTCTCTCCAGCTTCCCTCTCTCCACACACACCTTCTTCTCTCGCGCG 216
Qy 41 ProSerThrIleProThrSerProProProSerSerArgSerThrProSerAlaProPro 60
Db 217 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276
Qy 61 ProSerProProThrProSerThrProGlySerProProProProLeuProGlnProSerPro 80
Db 277 CCATCTCCACCAACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCCA 336
Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100
Db 337 CCCGTCCAACATACGCCCGGATCTCCACCGCACCTGTACTCTCTCTCTCTCTCTCTCT 396
Qy 101 ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro 120
Db 397 CCACCTTCAGTCCAGGACCACTCCATCTCTCAACGGAAGGAGGATCTCTCTCGACCT 456
Qy 121 ProSerProProSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140
Db 457 CCATCT 516
Qy 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160
Db 517 GGAATCGCATCGGAGGAGTCT 576
Qy 161 LysLysLysAtsArgArgAspGluAspAlaTyrTyrValProProProProProProPro 180
Db 577 AAGAGAAACGACGAGAGAGACGAGAGATGCTTACTATGTTCTCTCTCTCTCTCTCTCT 636
Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200

DB:	8	Gaps:	0
US-10-086-464-2 (1-647) x AY028699 (1-2189)			
QY	1	MetSerSerAlaProSerProGlyThrGlySerProGlySerProProSerProSerAsnSerThr	20
DB	97	ATGTCTCGCGCCGCTCTCCGGGAGCTGGTTCCGCTCCATCTCCACCATCAAACTCCACA	156
QY	21	ThrThrThrProProAlaSerAlaProProProProThrThrProSerSerProProPro	40
DB	157	ACCACACTCTCTCTCCAGCTTCGCGTCTCTCCACACACACCTTCTCTCTCCGCG	216
QY	41	ProSerThrIleProThrSerProProProSerSerArgSerThrProSerAlaProPro	60
DB	217	CCATCCACTATTCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	276
QY	61	ProSerProProThrProSerThrProGlySerProProProProProLeuProGlnProSerPro	80
DB	277	CCATCTCCACCAACTCCATCTACGCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCA	336
QY	81	ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro	100
DB	337	CCCGTCTCAACTACCGCCGATCTCCACCCGACCTGTTACTCTCTCTCTCTCTCTCTCTCT	396
QY	101	ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro	120
DB	397	CCACCTTCAGTCCAGGACACCGTCCAATCTTCCGCGAAGGAGGATCTCTCTCGACCT	456
QY	121	ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal	140
DB	457	CCATCT	516
QY	141	GlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys	160
DB	517	GGAAATCGCATCCGAGGAGTCTCTCTCTGTGTAGTAGTACTCTGATTGTCTCTCTCTGT	576
QY	161	LysLysLysArgArgAspGluGluAspAlaTyrTyrValProProProProProProPro	180
DB	577	AAGAAGAAACGACGAGAGACGAAGATGTTACTATGTTCTCTCTCTCTCTCTCTCTCTCT	636
QY	181	GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnThrArgGlnGlnAsnAla	200
DB	637	GGTCCCAAGCCGAGGACCTTACCGGTGACAGCAGCAACAATGGCGGCAACAAACGCA	696
QY	201	ThrProProSerAspHisValValThrSerLeuProProProProProProProProPro	220
DB	697	ACACCAACCGTCAGATCATGTCTGTAGCTACTTACCACCAACCACTTACGCTCCATCTCCA	756
QY	221	ProArgGlnProProProProProProProPheMetSerSerSerGlyGlySerAsp	240
DB	757	CCACGGCAACCTCTCTCCACCTCCACCGCCTTTTCATGAGCAGCAGCGCGCTCCGAC	816
QY	241	TyrSerAspArgProValLeuProProProProProGlyLeuValLeuGlyPheSerLys	260
DB	817	TACTCGGACCGTCCAGTTCTCTCTCCCGCTCTCCAGGCTTGTGTAGGCTTCTCCAAA	876
QY	261	SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu	280
DB	877	AGCACTTTTCATATCAGAGAGCTAGCTAGAGCCACAATGGTTTCTCCGAGCGCACTTG	936
QY	281	LeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal	300
DB	937	TTAGGACAAGCGGGTTCGGTTTACGTGCGACAAAGGTGTGTCTGCTAGTGGGAAAGATT	996
QY	301	AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal	320
DB	997	GCTGTGAAGCAGTTGAAGTTCCGGAGTGGTTCAGGAGAGAGAGGAGTTTCAGGACAGGTT	1056
QY	321	GluIleIleSerArgValHisHisArgHisLeuValSerLeuValClyTyrCysIleAla	340
DB	1057	GAGATCATCAGCAGAGTTCCACACAGGCATCTGGTGTCTCTTGTGGTTATTGCAATCGCC	1116
QY	341	GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuLeuHisLeu	360
DB	1117	GGTGCCAAAGATTGCTTGTCTATGAGTTGTCTTAACAACAATCTCGAGCTTCACCTC	1176
QY	361	HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer	380
DB	1177	CATGCGAGGACGCGCTACAAATGGAATGGAGACCAAGATTGAAGATTGCTCTTGGATCT	1236
QY	381	AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle	400
DB	1237	GCTAAAGGACCTTCTTATCTTCATGAAGATTGCAATCTTAAATCATTTCCCGTGATATC	1296
QY	401	LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu	420
DB	1297	AAGGCTTCAAAACATATTGATAGATTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT	1356
QY	421	AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly	440
DB	1357	GCTAAGATTGCTTCTGATACAAACACGCATGTATCAACACGTGTATGGAAACCTTTGGG	1416
QY	441	TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer	460
DB	1417	TACTTGGCTCCGAATACGCTGCAAGCGGAAAGCTCACGGAGAGCTCTGACGTTTCTCA	1476
QY	461	PheGlyValValLeuLeuLeuIleThrGlyArgArgProValAspAlaAsnVal	480
DB	1477	TTTGGCGTTGTGCTTTTGGAGCTCAATTACTGACGCTCGACCGCTTGATGCCAACATGTC	1536
QY	481	TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln	500
DB	1537	TATGTAGATGACAGCTTAGTTGACTGGGCACGACCAATGCTTTAACCGAGACTCTGAGCAA	1596
QY	501	GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluMet	520
DB	1597	GGAGACTTTTGGGTTTAGCTGATGCAAAAGATGAATATGGGTATGACAGAGGAGATG	1656
QY	521	AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg	540
DB	1657	GCTCGCATGTTGCTTGTGCTGCGCTTGTGTTGCCATTTCAGCTCGCGCAGACCTCGC	1716
QY	541	MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly	560
DB	1717	ATGACCCAGATTGTGCGTGGCTTAGAAGAAATGATATCAGTGTGAGTCTTAAACGAAGG	1776
QY	561	MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer	580
DB	1777	ATGACACCGGTCAAGCAATGTATACAGCTCATACGGAGGAAGCACCGATTATGACTCG	1836
QY	581	SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr	600
DB	1837	AGCCAGTACAATGAAGACATGAAGAAGTTTAGGAAAAATGGCACTTGGAACTCAAGAGTAC	1896
QY	601	AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer	620
DB	1897	AACGCCACGGGTGAGTACAGTAATCCGACCACTGACTATGAGCTGTACCCGCTCGTTCA	1956
QY	621	SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln	640
DB	1957	AGCAGCGAGGCCAACCCACACGCGAATGGAGATGGGGAAGATTAAAGAAACCGGTCA	2016
QY	641	GlyTyrSerGlyProSerLeu	647
DB	2017	GGTTATAGTGGACCTTCTCTT	2037
RESULT 5			
AX825738			
LOCUS	AX825738	1559 bp	DNA linear PAT 11-DEC-2003
DEFINITION	Sequence 36 from Patent WO03072763.		
ACCESSION	AX825738		
VERSION	AX825738.1 GI:39751255		
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 Goring, D., Silva, N. and Haffani, Y.Z.
Increasing plant seed production
Patent: WO 03072763-A 36 04-SEP-2003;
JOURNAL Goring, Daphne (CA); Silva, Nancy (CA); Haffani, Yosr, Z. (CA)

FEATURES

source

1. 1959
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Alignment Scores:

Pred. No.: 1.23e-63 Length: 1959
Score: 2936.50 Matches: 574
Percent Similarity: 90.61% Conservative: 24
Best Local Similarity: 86.97% Mismatches: 41
Query Match: 85.04% Indels: 21
DB: 6 Gaps: 12

US-10-086-464-2 (1-647) x AX825738 (1-1959)

```
QY 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17
DB 1 ATGTCCACAGCGCGTCTCCAGGCACACTCCATCACCATCTCCACCGTCTCCCTCCCA 60
QY 18 AsnSerThrThrThrProProAla-----SerAlaProProThrPro 35
DB 61 AACTCGAACACCACTCTCTCTCCAGAGCCTCTCTCTCTCTCCACCACTCTCT 120
QY 36 SerSerProProPro-----ProSerThrIleProThrSerProProSerSerArgSer 54
DB 121 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 55 ThrProSerAlaProProProSerProProThrProThrProGlySerProProPro 74
DB 181 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 228
QY 75 LeuProGlnProSerProAlaProThrThrThrProGlySerProAlaProValThr 94
DB 229 CTTCCTCAACCTTCCCTCTCCGCTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCT 285
QY 95 ProProThr-----ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113
DB 286 CCTCAAAACCTCGAAGCCCTCCATCTCTTAACCAAGACCAACCAACCACTCCCTCA 342
QY 114 GluGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131
DB 343 ---GGATCTACTCTAGAACCTCCATCAACCACTAAACCGTGGCGCGTCTGATCTTCC 399
QY 132 AspGlyLeuSerThrGlyValValValValValValValValValValValValVal 151
DB 400 GATGGATTGCTACCGAGTTGTGTAGGAATCCCAATTGGAGAGTCGCTATCTTCTGTT 459
QY 152 IleValThrLeuIleCysLeuLeuCysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
DB 460 ATACTGACTCTGATTGTTCTCTCTGTGAAGAAGAAACGAGAACGACGACGATGAA 519
QY 170 AspAlaThrValProProProProProProProProGlyProLysAlaGlyProThrGly 189
DB 520 GCTGCTTACTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 579
QY 190 GlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 208
DB 580 GGTCAACACACAGTATTGGCAACACACACACACACACACACACACACACACACAC 639
QY 209 ThrSerLeuProProProProProProProProProProProProProProProProPro 228
DB 640 ACGTCAATTG---CCACCACTTAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 696
QY 229 ProProPro-----PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247
```

```
DB 697 CCACCCACCATTCATGAGTAGCAGTGGTGGTCTGACTATTCGGAATCTTCGGTCTT 756
QY 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
DB 757 CTTCCACCATCTCCAGGCTTGTGTAGGCTTTCTAAAGACACTTTTCACTTATGAGGAG 816
QY 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly 287
DB 817 TTGTGAGAGCTACTAATGCTTCTCTGAGGCTAAATTTGTAGGACAAAGAGGCTTGGT 876
QY 288 TyrValHisGlyGlyValLeuProSerGlyGlyValAlaValValValValValVal 307
DB 877 TATGTGCATAAAGTATATTTGCTAGTGGGAAGAAGTTGCTGTGAACAGTTGAAAGCT 936
QY 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuLeuLeuLeu 327
DB 937 GGTAGTGGTCAGGAGAGAGAGAGTTTCAGGCTGAGGTTGAGATCATTCAGAGATTCAT 996
QY 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
DB 997 CACAGGCATTTGGTCTCTTATTTGGTTATTTATGTCGCGGTCTCAAGATTAATCTTGC 1056
QY 348 TyrGluPheValProAsnAsnLeuLeuLeuHisGlyGlyGlyGlyGlyGlyGlyGly 367
DB 1057 TATGAGTTTGTTCACAAACAATCTTTGAGTTTCACCTCCATGTAAGGACGCGCTACG 1116
QY 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
DB 1117 ATGGAATGAGTACTAGATTGAAGATGCTCTTGGATCTGCTAAAGGACTTTCATATCTT 1176
QY 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407
DB 1177 CATGAAGATTGCAATCCGAAATCATTCACCGTGATTAAGCGCTCAACATATGAT 1236
QY 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
DB 1237 GATTTCAAATTTGAAGCTAAGTTGCTGACTTTGGTCTTCCCAAGATGCTTCTGTATACA 1296
QY 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGlyTyrAla 447
DB 1297 AACACTCATGTAATACACCGGTGATGGGAACCTTTGGGTATTTGGCTCCGGAATAGCT 1356
QY 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
DB 1357 GCAAGTGGAAAGCTCACAGAAAAGCTGACGCTTTCTCAITTTGCGGTGTACTTTTGGAA 1416
QY 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487
DB 1417 CTTATTACTGGGAGCGCCCTGTGTATGCCAAATGCTCTATGTAGATGACAGCTTAGTT 1476
QY 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507
DB 1477 GACTGGGCACGACCATTTGCTTAACCGAGCATCTGAGGAAGAGGATTTTGGGGTTGGCT 1536
QY 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
DB 1537 GATTCAAAAGATGGTAAATGATGATGACAGAGAGAGATGCTCGCATGGTGTCTGGCT 1596
QY 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
DB 1597 GCGGCTGTGTTTCCCAATTCAGTCCGCGCAGACCTCCGATGAGCCAGATAGTAGCGGCG 1656
QY 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
DB 1657 TTAGAGGAATATGATGCTGTCTGATCTTAAAGAGGATGAGACCGGGTTCACAGCAAC 1716
QY 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
DB 1717 GTATACAGCTCATATGGAGGAAGCACAGACTATGACACGAGCAATCAACACGACGATG 1776
QY 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSer 607
```


Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
 Db 1777 ATAAAGTTTAGGAAATGGCTCTTGGAACTCAAGATACGCCACACCGCGAGTACAGT 1836

Qy 608 AsnProThrSerArgTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThr 627
 Db 1837 AATCAACCACTGACTACGAGTCTGCTGGTCAAGCAGTGAAGGTCAAGCCACA 1896

Qy 628 ArgGluMetGluMetGlyLysIleLeuArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
 Db 1897 CGAAGATGGAGTGGGAAGATTAGAAACCGGTCAAGGTATAGTGGACCTCTCTT 1956

RESULT 8
 AY056788 2116 bp mRNA linear PLN 08-OCT-2001
 LOCUS Arabidopsis thaliana AT3g24550/MOB24_8 mRNA, complete cds.
 DEFINITION AY056788
 ACCESSION AY056788.1 GI:15983764
 VERSION F11 CDNA.
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana

REFERENCE
 AUTHORS Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayaashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusea,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
 TITLE Arabidopsis cDNA clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2116)
 AUTHORS Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayaashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusea,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
 TITLE Direct Submission
 JOURNAL Submitted (11-SEP-2001) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayaashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.
 Location/Qualifiers
 1. .2116
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"

/chromosome="3"
 /clone="RAFLO9-19-J24 (R14362)"
 /ecotype="Columbia"
 1. .36
 37. .1995
 /note="putative protein kinase"
 /codon_start=1
 /product="AT3g24550/MOB24_8"
 /protein_id="AAL10479.1"
 /db_xref="GI:15983765"
 /translation="MSTAPSGPTTPSPSPPTNNTTTTPPAASPPPTTPSSPP
 PSPSTSTSPSSPLPPSPSPGSLTPPLPQSPSPAPITPSPSPPTTPSPSP
 PSNPQPPNTPSGSTPRTPNTAPSPSPSDSLGVVVGIALGVAILVILICL
 LCKKRRRHDDAAIYVPPPPGPKAGGPGQQYQQQWQGNASRSDHVVLSLP
 PKPSPRPKPPPPPPAFMSSGGSDYLDPLVLPSPGVLGFSKSTFYELSR
 TNGFSANLIGCAGVQRLVYVFNPNLLEFLHKGKRPTEWSTRLKIALGSAKLSYL
 HLVSLLGYCMAGVQRLVYVFNPNLLEFLHKGKRPTEWSTRLKIALGSAKLSYL
 HEDCNPKIHRDKASNLIDYFKAADPGLAKIASDNTNTHVSTVMGTFGVLAPE
 YAASGLKTEKSDVPSFGVVLLELITGRRPVDANNVYDMSLDVWARLELNRASEGDF
 EGLADSKMGNEYDREEMVAVCAACVRSARRRPMQSVRALEGNVLSLDNEG
 RPKHSNYSYSGSTDYDTSQYNDMDIKFRKMAIGTQYGTGTSYNSPTSDYGLYPSG
 SSSEGOATREMGMGKIKTKQGYSGL"
 1996. .2116

3'UTR
 ORIGIN

Alignment Scores:
 Pred. No.: 1,3e-63 Length: 2116
 Score: 2936.50 Matches: 574
 Percent Similarity: 90.61% Conservative: 24
 Best Local Similarity: 86.97% Mismatches: 41
 Query Match: 85.04% Indels: 21
 DB: 8 Gaps: 12

US-10-086-464-2 (1-647) x AY056788 (1-2116)

Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProSer 17
 Db 37 ATGTCCACACGGCGCTCTCCAGGCACCTACCTCCATCACCCTCTCCCTCCACCA 96
 Qy 18 AsnSerThrThrThrThrProProAla-----SerAlaProProThrThrPro 35
 Db 97 AACTCGACACACCACTCTCTCTCCAGAGCCTCTTCTCTCTCTCCACCACTCT 156
 Qy 36 SerSerProProPro-----ProSerThrIleProThrSerProProSerSerArgSer 54
 Db 157 TCTTCT 216
 Qy 55 ThrProSerAlaProProProSerProProThrProSerThrProGlySerProPro 74
 Db 217 CCT 264
 Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94
 Db 265 CTTCCTCAACTTCCCTCTCCGCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCT 321
 Qy 95 ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113
 Db 322 CCCTCAAAACCTCGAAGCCCTCCATCTCTCTAAACCAAGGACCAACCACTCTCT 378
 Qy 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131
 Db 379 ---GGATCTACTCTAGAACTCCATCAAACTAAACCTGCGCGCGCTCTGATTCTTC 435
 Qy 132 AspGlyLeuSerThrGlyValValGlyIleAlaIleGlyGlyValAlaLeuVal 151
 Db 436 GATGGATTGTCTACCGAGTGTGGTAGGAATGCCATTTGGAGGAGTGGCTATTCTTGT 495
 Qy 152 IleValThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
 Db 496 ATACTGACTCTGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 555
 Qy 170 AspAlaThrTyrValProProProProProProProProProProProProProPro 189

```
Db 556 GCTGCTTACTATGTTCTCTCTCTCCATCTGCTCCCAAGAGCTGGAGGACCTTACCGT 615
Qy 190 GlyGlnGlnGlnTrpArgGlnGlnAenAlaThrProProSerAsp--HisValVal 208
Db 616 GGTCAACAACTAGTATTGGCAACAACAAACCGGTCACGGCGTTCAGATAATCATGTAGTG 675
Qy 209 ThrSerLeuProProProProlysAlaProSerProProArgGlnProProProPro 228
Db 676 ACGTCATTG---CCACCACTTAAGCCTCAATCCACCAACGAAACCTCTCCGCCACCT 732
Qy 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247
Db 733 CCACCACGAGCATTCATGAGTAGCAGTGGTGGTCTGACTATTCGGATCTTCCGGTTCTT 792
Qy 248 ProProProSerProGlyValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
Db 793 CCTCCACCATCTCCAGGCGCTGTGTAGGCTTTCTTAAGGACACTTTTCACCTATGAGGAG 852
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAenLeuLeuGlyGlnGlyPheGly 287
Db 853 TTGTCGAGAGCTACTAATGGCTTCTCTGAGGCTAATTTGTTAGGACAAGGAGGTTTGGT 912
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Db 913 TATGTGCATAAAGGTATATTGCTAGTGGAAAGATGCTGTGTAACACAGTTGAAAGCT 972
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleSerArgValHis 327
Db 973 GGTAGTGGTCAGGGAGAGAGAGGTTTCAGGCTGAGGTTGAGATCATTAGCAGAGTTTCAT 1032
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
Db 1033 CACAGGCATTTGGTTCTTATTGTTATGTTATGTCGGCGGTGTTCAAAGATTACTTCTGC 1092
Qy 348 TyrGluPheValProAsnAsnLeuLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367
Db 1093 TATGAGTTTGTTCCTCAACCAACAACTCTGAGTTTCACTCCATGTTGTAAGGAGCGGCTACG 1152
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
Db 1153 ATGGAAATGGAGTACTAGATTTGAAGATTTGCTTTGATCTGCTTAAGGACTTTTCATATCTT 1212
Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIle 407
Db 1213 CATGAAGATTGCAATCCGAAATCATTCACCGTGATATTAAAGCGTCAAAACATAATTGATT 1272
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1273 GATTTCAAATTTGAAAGCTAAAGTTGCTGACTTTGGTCTTGCCAAAGATTGCTTCTGATACA 1332
Qy 428 AenThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db 1333 AACACTCATGTATCTACACCGGTGATGGAACTTTGGGTATTGGCTCCGGAATATGCT 1392
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1393 GCAAGTGGAAAGCTCACAGAAAGCTGACGTTTCTCATTTGGCGTTGTACTTTGGAA 1452
Qy 468 LeuIleThrGlyArgArgProValAspAlaAenValTyrValAspAspSerLeuVal 487
Db 1453 CTTATTACTGGGAGCGCCCTGTTATCGCAACAATGCTATGTAGATGACAGCTTAGTT 1512
Qy 488 AspTyrAlaAatgProLeuLeuAenArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 1513 GACTGGGACGACCAATTCGTTAACCGAGCATCTGAGGAAGAGATTTTGAGGTTTGGCT 1572
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1573 GATTCAAAGATGGGTAATGAGTATGACAGAGGAGATGGCTCGCATGGTTCCTTGGCGCT 1632
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db 1633 GCGGCTTGTGTTGGCAATTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTACGGGCG 1692
```

```
Qy 548 LeuGluGlyAenValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1693 TTAGAGGAATGATCGCTGCTGAATCTTAACGAGGATGAGACCGGGTCACAGCAAC 1752
Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1753 GTATACAGCTCATATGGAGGAAGCACAGACTATGACAGCAATACACGACGACATG 1812
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1813 ATAAAGTTTAGAAATAGGCTCTTGGAACTCAAGAATACGGCACACCGCGGAGTACAGT 1872
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGlyGlnThrThr 627
Db 1873 AATCCACCACTGACTACGGAATGACCGCTGCTGTTCAAGCAGTGAAGGTCAGGCCACA 1932
Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
Db 1933 CGAGAAATGGAGATGGGAAAGATTAAAGAAACCGGTCAAGGTTATAGTGGACCTCTCTT 1992
RESULT 9
AY059901 2188 bp mRNA linear PLN 04-NOV-2001
LOCUS Arabidopsis thaliana protein kinase-like protein (MOB24.13) mRNA,
DEFINITION complete cds.
ACCESSION AY059901 GI:16649062
VERSION AY059901.1
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 2188)
AUTHORS Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
Palin,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
JOURNAL Submitted (22-OCT-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT e-mail for correspondence: arab@sequence.stanford.edu
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Nguyen,M.,
Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,
Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.
Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.
FEATURES
source
1..2188
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="III"
/clone="RAFL09-06-A17"
/note="this clone is in pBluescript"
1..2188
/gene="MOB24.13"
gene
```


Qy	229	ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspA-gProValLeu	247
Db	726	CCACCACCAGCATTCATGAGTAGCAGTGGTGGTCTGACTATTCGGATCTTCGGGTCTT	785
Qy	248	ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu	267
Db	786	CCTCCACCATCTCCAGGCGTGTGTAGGCTTTTCTAAAGACACTTCACCTATTAGAGGAG	845
Qy	268	LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly	287
Db	846	TTGTTCGAGAGCTACTAATGGCTTCTCTGAGGCTAATTTGTAGCACAGGAGGGTTGGT	905
Qy	288	TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal	307
Db	906	TATCTGCATAAAGTATATTCCCTAGTGGGAAGAAGTTGCTGTGAACACAGTTGAAAGCT	965
Qy	308	GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHis	327
Db	966	GGTAGTGGTCAGGAGAGAGAGAGTTTCAGGCTGAGGTTGAGATCATTCAGCAGATTTCAT	1025
Qy	328	HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuVal	347
Db	1026	CACAGGCAATTTGGTTCTCTTATTGGTTATTGTATGGCGGTGTTCAAGATTACTTTGTC	1085
Qy	348	TyrGluPheValProAsnAsnAsnLeuGluLeuHisIleuHisGlyGluGlyArgProThr	367
Db	1086	TATGAGTTTGTTCAAACAAATCTTTGAGTTTCACCTCCATGTAAGGACGCGCTACG	1145
Qy	368	MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu	387
Db	1146	ATGGATGGAGTACTAGATTGGAAGATTGCTCTGGATCTGCTAAAGACTTTCATATCTT	1205
Qy	388	HisGluAspCysAsnProLysIleHisArgAspIleLysAlaSerAsnIleLeuIle	407
Db	1206	CATGAAGATTGCAATCCGAAATCATTCACCGTGATATTAAAGGGCTCAACATATTCATT	1265
Qy	408	AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr	427
Db	1266	GATTTCAAAATTTGAAGCTTAAGTTGGTGACTTTGGTCTTGGCCAAGATTGCTTCGTATCA	1325
Qy	428	AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla	447
Db	1326	AACACTCATGTATCTACACGGCTGATGGGAACCTTTGGGTATTTGGCTCCGGAATATGCT	1385
Qy	448	AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu	467
Db	1386	GCAAGTGGAAAGCTCACAGAAAAGCTGACGCTTTTCTCATTTGGCGTTGTACTTTTGGAA	1445
Qy	468	LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal	487
Db	1446	CTTATTACTGGAGGCGCCCTGTGTGATGCGAACAATGCTGATGTAGATGACAGCTTAGTT	1505
Qy	488	AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla	507
Db	1506	GACTGGGCACGCCATTGCTTAAACGAGCATCTCAGGAAGGAGATTTTGAGGGTTTGGCT	1565
Qy	508	AspAlaLysMetAsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAla	527
Db	1566	GATTCAAAGATGGGTATAGATATGACAGAGAGAGATGGCTCGCATGTTGCTTGGCT	1625
Qy	528	AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla	547
Db	1626	GCGGCTTGTGTTCCGCAATTCAGCTCGCCGAGACTCGCATGAGCCAGATGATACGGGCG	1685
Qy	548	LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn	567
Db	1686	TTAGAAGGAATGATCGCTGCTGATCTTAAACGAAGGATGACACGGGTCACAGCAAC	1745
Qy	568	ValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet	587
Db	1746	GTATACAGCTCATATGAGGAAGACACAGCTATCACACGACCAATACACACGACATG	1805

Qy	588	LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlySer	607
Db	1806	ATRAAGTTTATAGGAAATATGGCTCTTGGAACTCAAGAAATACGGCAACAACCGCGAGTACAGT	1865
Qy	608	AsnProThrSerSerPtyrGlyLeuTyrProSerGlySerSerGlyGlyGlnThrThr	627
Db	1866	AATCCAAACCACTGACTACGGACTGTACCCGTCGTGGTTCAAGCACTGAAGGTCAAGCCACA	1925
Qy	628	ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu	647
Db	1926	CGAGAAATGAGATGGGAAGATTAAGAAAAACCGTCAAGTTATAGTGGACCCCTCTCTT	1985
RESULT 11			
LOCUS	BT008409	1959 bp mRNA linear	PLN 15-MAY-2003
DEFINITION	Arabidopsis thaliana At3g24600 gene, complete cds.		
ACCESSION	BT008409		
VERSION	BT008409.1	GI:30725491	
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1959)		
	Cheuk, R., Chen, H., Kim, C.-J., Shinn, P., Bowser, L., Carninci, P., Dale, J. M., Hayashizaki, Y., Huan, V. W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.		
TITLE	Arabidopsis ORF clones		
JOURNAL	Unpublished		
REFERENCE			
AUTHORS	2 (bases 1 to 1959)		
	Cheuk, R., Chen, H., Kim, C.-J., Shinn, P., Bowser, L., Carninci, P., Dale, J. M., Hayashizaki, Y., Huan, V. W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-MAY-2003) Salk Institute Genomic Analysis Laboratory (SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
COMMENT			
	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
	The Salk, Stanford, PGEN (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.-J., Shinn, P., Bowser, L., Chan, M. M., Chang, C. M., Dale, J. M., Huan, V. W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Southwick, A., Tang, C. C., Toriumi, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Yuan, S., Davis, R. W., Theologis, A., and Ecker, J. R.		
FEATURES			
source	Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PIs.		
	Location/Qualifiers		
	1..1959		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="3"		
	/clones="UI5366"		
	/ecotype="Columbia"		

Db 746 AGTCATTG---CCACCACCTAAGCCTCCATCTCCACCACGAAACCTCTCCGCGACCT 802

Qy 229 ProProPro---PheMetSerSerGlyGlySerAspTyrSerAspArgProValLeu 247

Db 803 CCACCACGAGATTCTAGTAGCAGTGGTGGTCTGACTATTGCGATCTCCGCGTTCTT 862

Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267

Db 863 CCTCCACCATCTCCAGGCGCTGTGTAGGCTTTTCTAAAGCACCTTTTCACTTATGAGGAG 922

Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly 287

Db 923 TTGTCGAGAGCTACTAAATGGCTTCTCTGAGGCTAATTTGTTAGGCAAGAGGGGTTTGGT 982

Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307

Db 993 TATGTGCATTAAGGTATATTGCTTAGTGGAAAGAGTTGCTGTGAACAGTTGAAAGCT 1042

Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327

Db 1043 GGTAGTGCTCAGGAGAGAGAGTTTCAGGCTGAGGTTGAGATCATTTAGCAGAGTTTCAT 1102

Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347

Db 1103 CACAGGCAATTTGGTTTCTCTATTGGTTATTGTATGGCGCGTGTTCAAAGATTACTTGTG 1162

Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlyValArgProThr 367

Db 1163 TATGATTTGTTTCCAAACACATCTTGAGTTTCACTTCCATGTGTAGGGAGCGGCTACG 1222

Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387

Db 1223 ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGAGTCTGCTAAGAGGACTTTTCATATCTT 1282

Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIle 407

Db 1283 CATGAAGATTGCAATCCGAAATCATTCACCGTGATATTAAAGCGGTCAAAATATTGATT 1342

Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427

Db 1343 GATTTCAAATTTGAAGCTTAAGTTGCTGACTTTGGTCTTGCCAAAGATTGCTTCTGTATCA 1402

Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447

Db 1403 AACACTCATGTATCATACCGGTGATGGAAACCTTTGGGTATTTGGCTCCGGAATATGCT 1462

Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467

Db 1463 GCAAGTGAAGAGCTCACAGAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTTGGAA 1522

Qy 468 LeuIleThrGlyArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487

Db 1523 CTTATTACTGGGAGCGCGCTGTGATGGCAACATGCTATGTAGATGACAGCTTAGTT 1582

Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507

Db 1593 GACTGGGCAGCACCATTGCTTAACCGAGCATCTGAGGAAGAGAGATTTTGGAGGTTTGGCT 1642

Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAla 527

Db 1643 GATTCAAAGATGGGTAATGAGTATCACAGAGAGGAGATGGCTCGCATGTTGCTTGGCCT 1702

Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547

Db 1703 GCGGCTTGTGTCGCATTCAGCTCGCGCAGACCTCGCATGAGGCACATAGTACGGGG 1762

Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567

Db 1763 TTAGAAGGAAATGATCGCTGTCTGATCTTAACGAAGGGATGAGACCGGGTCACAGCAAC 1822

Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587

Db 1823 GTATACAGCTCATATGGAGGAAGCAGACTATGACACGAGGCCAATACAAACGACACATG 1882

Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607

Db 1983 ATAAGTTTAGAAAATGGCTCTTTGGAACTCAAGAATACGCACACCGCGGAGTACAGT 1942

Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGlyGlnThrThr 627

Db 1943 AATCCAAACAGTACTACGGAAGTGTACCCGCTGTGGTTCAAGCAGTGAAGGTCAAGCCCA 2002

Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlnGlyTyrSerGlyProSerLeu 647

Db 2003 CGAGAAATGGAGATGGGAAAGATTAAAGAAACCCGCTCAAGGTTATAGTGGACCTCTCTT 2062

RESULT 13

AV089024 2324 bp mRNA linear PLN 14-APR-2003

LOCUS Arabidopsis thaliana clone 17909 mRNA, complete sequence.

DEFINITION Arabidopsis thaliana (thale cress)

ACCESSION AY089024

VERSION AY089024.1 GI:21407798

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 2324)

AUTHORS Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.

TITLE Full-length messenger RNA sequences greatly improve genome annotation

JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)

MEDLINE 22088475

PUBMED 12093376

REFERENCE 2 (bases 1 to 2324)

AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.

TITLE Full-length cDNA from Arabidopsis thaliana

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 2324)

AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.

TITLE Direct Submission

JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA

COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

source

1..2324

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="17909"

ORIGIN

Alignment Scores:

Pred. No.: 2,45e-62 Length: 2324

Score: 2885.00 Matches: 570

Percent Similarity: 90.30% Conservative: 26

Best Local Similarity: 86.36% Mismatches: 42

Query Match: 83.55% Indels: 23
DB: 8 Gaps: 13
US-10-086-464-2 (1-647) x AY089024 (1-2324)

QY 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProSer 17
DB 105 ATGTCCACAGCGCGCTCCAGGCACACTCCATCACCATCTCCACCGTCTCCGCCACA 164
QY 18 AsnSerThrThrThrProProAla-----SerAlaProProProThrPro 35
DB 165 AACTCGACACCACTCTCTCTCCAGGAGCTCTCTCTCTCCACCACTCTCTCT 224
QY 36 SerSerProProPro-----ProSerThrThrThrProProProSerSerArgSer 54
DB 225 TCTTCT 284
QY 55 ThrProSerAlaProProProSerProProThrProSerThrProGlySerProPro 74
DB 285 CCT 332
QY 75 LeuProGlnProProProAlaProThrThrProGlySerProProAlaProValThr 94
DB 333 CTCTCTCAACCT 389
QY 95 ProProThr-----ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113
DB 390 CCTCTCAAAACCTCTCAAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 446
QY 114 GluGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131
DB 447 ---GGATCT 503
QY 132 AspGlyLeuSerThrGlyValValValGlyValAlaIleGlyGlyValAlaLeuVal 151
DB 504 GATGATGTCTACCGGAGTGTGTGTAGTAATCGCCATTGGAGAGTCTGCTATCTTGT 563
QY 152 IleValThrLeuIleCysLeuLeuCysGlyLysLysArgArg-----AspGluGlu 169
DB 564 ATACTGCTCTGATTTGCTCTCTGTAAGAAGAAAGCAAGAAAGACGACGATGAA 623
QY 170 AspAlaTyrValProProProProProGlyProGlyProLysAlaGlyGlyProTyrGly 189
DB 624 GCTCTTACTATGTT-----CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 680
QY 190 GlyGlnGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAsp-----HisValVal 208
DB 681 GGTCAACACAGTATTGGCAACACAAACCGCGTCACGGCGTCAGATAATCATGTAGT 740
QY 209 ThrSerLeuProProProProProLysAlaProSerProProArgGlnProProProPro 228
DB 741 ACGTCAATG---CCACCACCTTAAGCCTCCATCTCCACACGAAACCTCTCTCCGCCACCT 797
QY 229 ProProPro-----PheMetSerSerSerGlySerAspTyrSerAspArgProValLeu 247
DB 798 CCACCACAGCATTCATGAGTAGCAGTGGTGGTTCGACTATTCGGAATCTTCGGTTCCT 857
QY 248 ProProProProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
DB 858 CCTCCACCACTCCAGGCTTGTGTAGGCTTTTCAAAAAGCACTTTCATCTATGAGGAG 917
QY 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
DB 918 TTGTCAAGAGACTACTAATGGCTCTCTCTGAGGCTAATTTGTTAGGACAAGGAGGTTTGGT 977
QY 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
DB 978 TATGTGCATAAAGGTATATTGCTAGTGGGAAAGAGTTGCTGGAACAGTTGGAAGCT 1037
QY 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLeuSerArgValHis 327
DB 1038 GGGAGTGGTCAGGGAGAGAGAGATTTTCAGGCTGAGGTTGAGATCATTAGCAGAGTTTCAT 1097

QY 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuVal 347
DB 1098 CACAGGCATTTGGTTTCTCTTATTTGGTTATTTGGCTGATGTTCAAGATTTGCTTGT 1157
QY 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367
DB 1158 TATGAGTTTGTCTTAAACAATCTTTGAGTTTCACTTCCATGTTAAGGAGGAGCCCTACA 1217
QY 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
DB 1218 ATGGAATGGAGTACTAGATTGAAGATGCTTGGATCTGCTAAAGGACCTTTTCATATCTT 1277
QY 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407
DB 1278 CATGAAGATTGCAATCCGAAATCATTCACCGTGATATTAAAGCGCTCAAAACATATTGATT 1337
QY 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
DB 1338 GATTTCAAATTTGAGCTAAGTTTGTGACTTTTGGTCTTCCAGGATGCTTCTTGATACA 1397
QY 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
DB 1398 AACACTCATGTATCTACACCGCTGATGGAAACCTTTGGGTATTTGGCTCCAGAATATGCT 1457
QY 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
DB 1458 GCAAGCGGAAGCTCACAGAAAAGTCTGACGTTTTCTCATTTGGTGTGTACTTTTGGAG 1517
QY 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
DB 1518 CTTATATACCGGAGGCGCCCTGTTGTGCAACAACTGCTATGTAGATGACAGCTTAGTT 1577
QY 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507
DB 1578 GACTGGSCACGACCTTGTCTTAACCGAGCATCTGAGGAGGAGATTTTGTGTTGGCT 1637
QY 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
DB 1638 GATTCAAGATGGTATAGTATAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 1696
QY 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
DB 1697 GCAGTTTGTGTTCCCATTCAGCTCCGCCAGACCTCGCATGACCCAGATAGTAGGCGG 1756
QY 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
DB 1757 TTAGAAGAAATGTATCGCTGCTGATCTTAAACGAAGGGATGAGACCGGGTCACAGCAAC 1816
QY 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
DB 1817 GTATACAGCTCATATGGAGGAGCACTGACTATGACACGAGCCAAATCAACAGATGACATG 1876
QY 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSer 607
DB 1877 AGAAGTTTAGGAAATGGCTCTTGGAACTCAAGATACGGCACACCGCGGAGTACAGT 1936
QY 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGlyGlyGlnThrThr 627
DB 1937 AATCCACCACTGACTACCGGACTGTACCCGCTGCTTCAAGCAGTGAAGGTCAAGCCACA 1996
QY 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
DB 1997 CGAATAATGGAGATGGGAAAGATTAAGAAAAACCGGTCAAGGTTATAGTGGACCTCTCT 2056

RESULT 14
AB020746
LOCUS
DEFINITION
AB020746 BAO00014
ACCESSION
VERSION
AB020746.1 GI:3985949
KEYWORDS
SOURCE
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (sites)
 Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S. Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC and BAC clones
 DNA Res. 7 (3), 217-221 (2000)

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

2 (bases 1 to 79706)
 Kaneko,T., Katoh,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S. Direct Submission
 Submitted (03-DEC-1998) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kikarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)
 Address for correspondence: kaos@kazusa.or.jp
 For the latest information on annotation of this clone, please see <http://www.kazusa.or.jp/kaos/cgi-bin/agd/graph.cgi?c=MOB24>
 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <http://compbio.ornl.gov/Grail-1.3/>), GENSCAN (Chris Burge, MIT, <http://CCP-081.mit.edu/GENSCAN.html>), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://grelini.zool.iastate.edu/cgi-bin/sp.cgi>).
 Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).
 This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions.
 The 5' clone is MXP5 and the 3' clone is MSD24.

FEATURES

source

1..79706
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="MOB24"
 /clone_lib="Mitsui P1"
 2398..3399
 /note="gene_id:MOB24.1
 unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB01995.1"
 /db_xref="GI:9294038"
 /translation="MGDSEDETGYPKPKFYSILNRQNHMYSRPIPKRHAYYNEEDEDE
 VEGEEDPQGGYIRNGERFQKQKNVVSFGFEPAGPSDAKAVDWRQEAFLVLEW
 GDRFQLGRSLRDNWEAEKSEELMEKSETQCRMTDILKRYKKEKIKVEKS
 GLGSSKFFNKLMLLVCSPKSLGLACGVDGSGFVFMMTKVYLDKSNGLDMDSP
 GDSSEEDDEVEYERKYNDAASYKMLADSVRFKGYEKMEKSKQKLEKMR
 ADPQDLQKQIVDRAQSEIARUREEENHHGGDDDESEEMENDSDVNLSDS"
 complement(join(3794..4276,4363..4490,4577..4673,
 4757..5179,5354..5548,5693..5970,6088..6214,6297..6496,
 6601..6648,6743..6776,6882..7013,7464..7559,7677..7743,
 8162..8217,8290..8385,8470..8998,9122..9462))
 /note="gene_id:MOB24.2"
 /codon_start=1
 /evidence=not experimental
 /product="DNA repair protein MutS"
 /protein_id="BAB01996.1"
 /db_xref="GI:9294039"
 /translation="WQRQSRILSFQKPTAATKGLVSDAASGGGSGGPRFNVKEG
 DAKGDSVRFAVSKSVDVGRDTPTEKVPFRVLSPGPKPAESGASDASSLSFNHKKF
 VKVDRDCSGERSREDVPLNDLSLCMKANDVIFQFRNNGKQIQRNHAFFSFGRAEL

CDS

gene_id:MOB24.5
 similar to unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB01999.1"
 /db_xref="GI:9294042"
 /translation="WPLISASPPWPLSLSPSSSSLNRAAPSIPIPSFSPRINHGF
 SCSTIAVQRFALSKRLQQLVAMAPPEEKLTRNPLDFFIEWRPPRRPDIFPK
 FPMKSPPLPPMPDPPPEEEEEKKEEEDDDPEKEDQPDNRS"
 join(23402..23459,23819..24021)
 /note="gene_id:MOB24.6
 unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB02000.1"
 /db_xref="GI:9294043"
 /translation="WKNASLKLPLLIIFLITVLSNLGAERKLTGVADVIVEGAASAPQ
 YGIDENDTLNKSCKRDIQCTQCRGGFCNLLOKCECL"
 complement(join(25945..26640,26714..26824,26914..27012,
 27119..27224,27291..27360,27436..27583))
 /note="Contains similarity to ubiquitin conjugating enzyme
 gene_id:MOB24.7"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB02001.1"

CDS

gene_id:MOB24.7
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB02001.1"

CDS

gene_id:MOB24.7
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB02001.1"

RSVEDIGVDGVPGPETGMPRASRLKRVLEDEMTFKEDKVPVLDSNKRKLMLODPV
 CGEKKYNEGTKEFWLESSIRIDANRRRPPDLYDRKTLHIPPPVFKKMSQKQYWS
 VSKYMDIVLFPFKVEVELYELDAELGHEKLDWKMTSGKQVQKQVIGSSGIDEAV
 KKLARGVKYGRIFQLETSDOAKARGANTIIIPKLVOLVLTPTSTASEGNIGPDVAHLLA
 IKETIKLCKSTYVGFVDCALRFVWGSISDDASCAALGALLMOVSPKEVLVDSK
 GLSREAKRLKTYLTGSTVQLAPVQVMGDTDAAGVRNIIIESNGYFKGSSSEWNCA
 VDLNCEQVALSALGINHLRLKLEJLVKHGDIFFQVYRGRLRIDGQTMVNLEIF
 NNSCDGPGSTLYRYLDNCVSPTRKLRNWIHPLKQVDESINKRLDVFEFTANSES
 MOITGOYLHLKLPDLRLLGRKSSVRSSASVLPALLGKVLQKQKAFQVQVKGFRSG
 IDLLALQKESNMMLLYKLKLPILYKSGLELFLSQFEAAIDSDRPNQONQVDTDE
 NAETLILIELPIERATQWSEVIHTISCLDLVRSFATAASLSAGSMARPVIPPESEAT
 DQNTKGPILIKIQLWHFPAADGQLPVPNDILLGEARKSSGSIHRSLLLTGPNM
 GSKTLRATCLAVIFAQGCYPCSEISLVDITFRLGASDRNITGEGSTFLVECT
 ETASVLNATQDSILVIDELGRGTSFTDGAIAYSVFRHLVEKQCRMLFATHVPLT
 KEFASHPRVTSKHMACAFKSRSDYQPRGCDQDLVFLRLTEGACPEVGLQVAMAGI
 PNQVETASGAQAQMKSGISGENFKSSELSSEFSSLHEDWLKSLVGISRVAHNAPIGE
 DDYTLFCLWHEIKSSVCVPK"
 10131..10577
 /note="gene_id:MOB24.3"
 /codon_start=1
 /evidence=not experimental
 /product="ethylene-responsive transcriptional
 coactivator-like protein"
 /protein_id="BAB01997.1"
 /db_xref="GI:9294040"
 /translation="MPSRYPGAVTQDWPEPVVLHKSQKQSLDRDPKAVNAALRNGAV
 QTVKFPDAGSNKKGSTAVPVINTKLEETEPEAPAMDRVKAERLMIQKARLEKMSQ
 ADLAKQINERTQVQVEYENGKAVNOAVLAMEKVLGVKLRGKIGK"
 complement(join(11108..11381,11467..11528,11699..11836,
 11934..12071,12161..12334,12421..12510,12718..13104,
 13623..13765,14306..14405))
 /note="gene_id:MOB24.4"
 /codon_start=1
 /evidence=not experimental
 /product="aldehyde dehydrogenase"
 /protein_id="BAB01998.1"
 /db_xref="GI:9294041"
 /translation="WENKCNCGATTVKLPEIKTKLFGINGQIFDAASQKTFETIDPRN
 GEVIATAGDKEDVDLANAARYAFQHPMPRTMGTGERAKLINFKADLIENIEELA
 KLDAVGGKLFQGLKYADIPTAGHFRYAGAAADKIHGTELTMRQSLQVYHLKSKEAGIDPVL
 VVGNLIIPNPFSPIMFATKVPAMAAGCTMVVKPAEQTSLSALFYAHLKSKEAGIDPVL
 NIVTGFSTAGAAIASHMDVKVSTGSDTVGRKIMQAAASNLKVSLELGGKSPILL
 IFNDADIDKAADLALLGCFYNGKEICVASSRVFQEGYIDKVRKLEKAKDWTVGDP
 FDSARQGPQVDRQEFKILSYIEHKNEGATLLTGGKAIQDGKGYFIQFTIFADVTE
 MKIYQDRIFGPMVSLMKPFVEEGIKCANMTKYGLAAGILSQDILINTVSRISKAGI
 IWVNYGDFGLDCPYGKYMCGNCRSGMDALDNVLTQKSVVMPVHNSPMW"
 complement(join(19742..19933,20400..20657))
 /note="gb|AAD25142.1
 gene_id:MOB24.5
 similar to unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB01999.1"
 /db_xref="GI:9294042"
 /translation="WPLISASPPWPLSLSPSSSSLNRAAPSIPIPSFSPRINHGF
 SCSTIAVQRFALSKRLQQLVAMAPPEEKLTRNPLDFFIEWRPPRRPDIFPK
 FPMKSPPLPPMPDPPPEEEEEKKEEEDDDPEKEDQPDNRS"
 join(23402..23459,23819..24021)
 /note="gene_id:MOB24.6
 unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB02000.1"
 /db_xref="GI:9294043"
 /translation="WKNASLKLPLLIIFLITVLSNLGAERKLTGVADVIVEGAASAPQ
 YGIDENDTLNKSCKRDIQCTQCRGGFCNLLOKCECL"
 complement(join(25945..26640,26714..26824,26914..27012,
 27119..27224,27291..27360,27436..27583))
 /note="Contains similarity to ubiquitin conjugating enzyme
 gene_id:MOB24.7"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB02001.1"

CDS

CDS

CDS

CDS

CDS

/db_xref="GI:9294044"
/translation="MAQAARSLRMQKELKLLSDPPHGAASFPHLSAANGSGDPTF
STIDAQEGBEDTVYANGFNLIKQIPERYFPQPIVSFATPIVHPINDNSGICLDI
LNPYPGAWQPSLNI STVLSMRLLSGSDGDMGVSEKYNKRYNTQTFYKAREM
EKYAKVADGQSLTQIKHNGEKSGESNGSVYKLT VESSLSYKHTQTFYKAREM
EDNGKR KAVVFGEGNSFGNDTQIKRSKRLSLALPSQSKDLCGSEELTRGVSAAC
KENKNPLNGLKISLGLQFGNDTILAFRTSAAKSDNNRLSKSLGVLELVNYSK
PEVLATQDMKLENNQNDARSILRFGFENSVL EETSMABSI VLVLSDSDSGQEEERVSS
SRSLSLAKRVLVKCRP"

CD5

/note="emb|CAB79678.1
 gene_id:MOB24.8
 similar to unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB02002.1"
 /db_xref="GI:9294045"
 /translation="MSSLIAYLILSFNSFSLPKLRISLNLNLVLRSTWNIKKRWCL
 AIPYLHLHATPLWKSQITLMALRFSQMVPEPTVYFELSLSPAFLLTQAMSHGETLWD
 TTSLEGKTRLAFSSPSSPHGQNKQSPFSPENDASLSSPYNSRHHDLHLISEGL
 GLQVLGVFLGFLRYPVRSKQDKQLPKLRSLDANVLSLIPFSSQSLHYLLASLDIRGHH
 TNNELVLIWGLRNLNRTPLPSKRVFVFGILKNGNIMGIPSTLRWCLDLCRPSAT
 CISPTHLGISFKQCSLSRRLKAKIRISIKRSEKVNPLFEIETESPHALVWGLKGLGIMI
 TSPRSQDQRFNLSHSYPLTEFKQKRNLIKFSIHQSLGKNGIMLSSPQMSDYYR
 SPFNLDLDTPTVHLLGDLQGLQLLVASQSALEPYSTSRLLTNTVNLISIDSFYED
 PSTTRDITCAIMPEYSLWLKALANFMSFHIYIIVHLVGNASNYVYVLENFVESLLLI
 FV
 Li
 101n(32831..33058,33143..33907)

CDS

```

factor
  gene_id: MOB24.9
  /codon_start=1
  /evidence=not experimental
  /protein_id="BAB02003.1"
  /db_xref="GI:9294046"
  /translation="MEDDNNNNNNNNVIPIVKTQVNDPSTDMLITWGAHNSF
IVDPDLSRIQLPFIKNNFSFVRQLTQVIFGPKVDQDKEWFANFHELRQGHLLN
NIARHARGMYQGLDEGIEVREIRELKEQRELEAEIQMRNRRIEATSKPEQWMA
FLYKVEDPDLPLNNKERTQKQVQDKKKRVTVTSKSEESVEDEGVRVY
MSSSTPSPSTENLYRNHSPDGMVPMQTQVFGSVEGLVAKMSLNSTSGSSLT
TFLSPESVGGGGCGGSIQGRRYVKETATFGVVESNPPTTPYPFSLRFGGF"
  complement(join(317054..37590,37670..37809,37894..38197,
38288..38398,38486..38584,38662..38772,39015..39066,
39311..39367))
  /notes="gene_id:MOB24.10
unknown protein"
/codon_start=1

```

CDS

```

38388..38398,38486..38584,38662..38772,39019..39066,
39311..39367))
/notes="gene_id:MOB24.10
unknown protein"
/codon_start=1

7.29e-55          79706
Length:
Matches:          573
Conservative:    23
Mismatch:         43
Indels:          337
Gaps:             19

17) x AB020745 (1-79706)

```

IIS-10-08

QY	1	MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17
		::::
Db	51787	ATGTCCACAGGGCGTCTCCAGGCACACTCCCATCACCAATCCTCCACGGTCCTCTCCACCA 51846
QY	18	AsnSerThrThrThrProProproAla-----SerAlaProProProThrThrPro 35
Db	51847	AACTCGACAACCACCACTCTCTCTCCAGAGGCTCTTCTCTCTCTCCACCAACAACCTCT 51906
QY	36	SerSerProProPro---ProSerThrIlePLeProThrSerProProProSerSerArgSer 54
Db	51907	TCTTTCTCTCTCCGTCGGCGTCAACTAATTCAACCTCTCTCTCTCTCTCTCTCTCTCTCTTA 51966
QY	55	ThrProSerAlaProProProSerProProThrProThrProSerProGlySerProProPro 74
Db	51967	CCT 52014
		-----CCTGCT 52014

Fujimura, T., Fukuda, S., Haneagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Kikuchi, Hori, F., Hotta, I., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Koijima, K., Koijima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.>

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Tehibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusunegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niihara, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hata, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hikamoto, K., Hixao, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kiehkawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishik, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

```

1. .2307
   Location/Qualifiers
     organism="Oryza sativa (japonica cultivar-group)"
     mol_type="mRNA"
     cultivar="Nipponbare"
     db_xref="taxon:39947"
     clone="J033121K23"

```

[illegible]

US-10-086-464-2 (1-647) x AK103247 (1-2307)

[illegible]

Job time : 5988 secs

```
Db 988 TATGTTTCAAAAGGAGTTCTGCGAATGGGACAGAGGTTGCTGTGAAGCAATTGAGAGAT 1047
QY 308 GlySerGlyGlnGlnGlyGlnArgGlnGlnValGlnValGlnValGlnValGlnValHis 327
Db 1048 GGGAGTGGGCGAGGAGAGCGGTAATTTTCAGCGGAGGTTGAGATTATCAGCGGGTACAT 1107
QY 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaGlyAlaGlyLeuVal 347
Db 1108 CACAAGCATCTGTGAACATTTGGTTGGTTATGTCATTTCTGGAGGAAGAGGTTGCTTGTC 1167
QY 348 TyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367
Db 1168 TATGAGTATGTTCCAAAACAACATTTGGAGCTCCACTTGCATGGAGAGCGCCACCAACC 1227
QY 368 MetGluTyrPheThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
Db 1228 ATGGAATGGCTTACAAAGATTACGCATTTGGGTGCTGCGAAGGGATTGGCATATCTT 1287
QY 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIle 407
Db 1288 CATGAAGATTGGCATCTTAAGATCATTCACGTGACATAAAGTCAGCAAAATATTTCTT 1347
QY 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1348 GACGGAGAGATTGAGGCAAAAGGTGGCAGATTTCGGACTTGCAAAATTGACTTCTGATAAT 1407
QY 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db 1408 AACACTCATGTTTCAACAGAGTAAATGGGCATTTGGGTACCTTGGCACCAGAGTATCG 1467
QY 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1468 TCTTCTGGCAGCTAACTGAGAAATCAGATGCTTTTTCTTTCGGAGTAATGCTTCTTGAG 1527
QY 468 LeuIleThrGlyArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487
Db 1528 CTAATAACTGGCGCGCGTCTGTGAGATCGAACCAATCACAGATGCATGCATGGTGGTT 1587
QY 488 AspTyrAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 1588 GATTGGGAAGGCCCTTTAATGATGGCAGCATCAGATGATGGCAATTATGATGCTTTAGTG 1647
QY 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1648 GATCCACGGCTAGGACAGGAGTACAATGGTAACGAGATGGCGAGGATGATTGCTTGCA 1707
QY 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db 1708 GCTGCTTGTGTACGCCATTTCTCAGCGCGACGCCACCGGATGAGCCAGGTTGTCGGGCC 1767
QY 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGlyMetArgProGlyGlnSerAsn 567
Db 1768 TTGGAAGCGGATGTGTCATTGGATGATCTGAACGAAGGTGTTGCTCGGCCATAGCCGA 1827
QY 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1828 TTCTTGGGATCATAC---NACAGCAACGAGTACGACACGGGCCATTACAAATGAAGACCTG 1884
QY 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1885 AAGAAATTTCAGGAAGATGGCATTTGGCAGTGGCAATCTA-----GAGAGCAGC 1932
QY 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerSerGluGlnThrThr 627
Db 1933 CAGCAAAACGCCACCCAGGTTGCTCCGGAATCGATCCGTATCAATGGGGGATGCGCGT 1992
QY 628 Arg-----GluMetGluMetGlyLysIleLysArgThrGly 639
Db 1993 CAATACCGGACGAGATGGAGATGGGAGTGGGAGTTTGAGAAAGATGGC 2040
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 24, 2004, 18:28:17 ; Search time 604 Seconds
(without alignments)

4550.636 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSAPSPGTCSPSPSPNST.....REMEMKIKRTGGYSGPSL 647

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10086464/runat_23042004_083014_5603/app.query.fasta_1.839
-DB=N-Geneseq_29Jan04 -CFM=fastap -SUFFIX=ring -MINMATCH=0 -1 -LOOPCL=0
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=txt -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10086464 @CGN 1 1 708 @runat_23042004_083014_5603 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3453	100.0	1944	4 AAF77094	Aaf77094 Brassica
2	3453	100.0	1944	9 ACF36548	Acf36548 B. napus
3	3453	100.0	2189	9 ACF36549	Acf36549 B. napus
4	2936.5	85.0	1959	9 ACF36557	Acf36557 A. thalia
5	2885	83.6	2324	3 AAC36968	Aac36968 Arabidops
6	1821.5	52.8	1812	9 ACF36554	Acf36554 O. sativa
7	1668	48.3	1515	9 ACF36555	Acf36555 Z. mays P
8	1637.5	47.4	1488	9 ACF36556	Acf36556 G. max PE

9	1507	43.6	2520	4 AAF77096	Aaf77096 Arabidops
10	1504	43.6	2520	9 ACF36551	Acf36551 Arabidops
11	1479.5	42.8	2820	4 AAF77097	Aaf77097 Arabidops
12	1471.5	42.6	2820	9 ACF36552	Acf36552 Arabidops
13	1209.5	35.0	2880	4 AAF77095	Aaf77095 Arabidops
14	1209.5	35.0	2880	9 ACF36550	Acf36550 Arabidops
15	1185	34.3	3060	9 ACF36553	Acf36553 Arabidops
16	1181	34.2	3060	4 AAF77098	Aaf77098 Arabidops
17	839.5	24.3	1923	7 ADA70890	Ada70890 Rice gene
18	816.5	23.6	2124	6 ABZ13274	Abz13274 Arabidops
19	804.5	23.3	2865	7 ADA70701	Ada70701 Rice gene
20	802.5	23.2	2011	4 AAH22557	Aah22557 P. patens
21	797	23.1	1730	7 ADA70647	Ada70647 Rice gene
22	793.5	23.0	2823	7 ADA69478	Ada69478 Rice gene
23	782.5	22.7	1419	3 AAC43029	Aac43029 Arabidops
24	782.5	22.7	1419	6 ABZ13121	Abz13121 Arabidops
25	775.5	22.5	1626	3 AAC48313	Aac48313 Arabidops
26	775.5	22.5	1917	7 ADA70527	Ada70527 Rice gene
27	773	22.4	1479	3 AAC43040	Aac43040 Arabidops
28	769.5	22.3	1353	3 AAC42860	Aac42860 Arabidops
29	769.5	22.3	1353	6 ABZ14106	Abz14106 Arabidops
30	766	22.2	1494	3 AAC48560	Aac48560 Arabidops
31	758.5	22.0	1572	3 AAC42898	Aac42898 Arabidops
32	758.5	22.0	1572	6 ABZ13705	Abz13705 Arabidops
33	758.5	22.0	1920	6 AAL45905	Aal45905 A thalian
34	758.5	22.0	2012	3 AAC44847	Aac44847 Arabidops
35	756	21.9	3239	3 AAA79271	Aaa79271 Eucalyptu
36	754.5	21.9	2136	7 ADA69829	Ada69829 Rice gene
37	749.5	21.7	1751	3 AAC45313	Aac45313 Arabidops
38	744.5	21.6	1485	3 AAC42702	Aac42702 Arabidops
39	740.5	21.4	3525	7 ADA71170	Ada71170 Rice gene
40	740	21.4	2868	3 AAA79266	Aaa79266 Pinus rad
41	740	21.4	3153	7 ADA71066	Ada71066 Rice gene
42	736.5	21.3	1703	3 AAC46629	Aac46629 Zea mays
43	736	21.3	3624	7 ADA69512	Ada69512 Rice gene
44	735	21.3	2004	3 AAC40780	Aac40780 Arabidops
45	734	21.3	1455	3 AAC45877	Aac45877 Arabidops

ALIGNMENTS

RESULT 1
AAF77094
ID AAF77094 standard; DNA; 1944 BP.
XX
XX
XX AAF77094;
DT 17-MAY-2001 (first entry)
XX
XX Brassica napus PERK1 DNA.
XX
XX Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.
XX
XX Brassica napus.
XX
XX WO200114563-A1.
XX
XX
XX PD 01-MAR-2001.
XX
XX PF 18-AUG-2000; 2000WO-CA000966.
XX
XX PR 19-AUG-1999; 99US-0149466P.
XX
XX PR 13-OCT-1999; 99US-0159122P.
XX
XX (GORI/) GORING D.
XX
XX (SILV/) SILVA N.
XX
XX Goring D, Silva N;
XX
XX WPI; 2001-244305/25.
XX
XX PT New proline-rich, extensin-like receptor kinase nucleic acids and
XX
XX PT polypeptides useful for increasing plant wounding or pathogen resistance,

or for producing transgenic plants with increased wounding or pathogen resistance.

Claim 6; Fig 1; 9lpp; English.

The present invention relates to proline-rich extensin-like receptor kinase (PERK). The PERK nucleic acids and polypeptides are useful for increasing the resistance of plants to wounding and pathogens. These are also useful for producing transgenic plants with increased wounding and pathogen resistance compared with a wild type plant, as well as in assays for identifying and developing compounds to inhibit and/or enhance polypeptide function directly

Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,628-93 Length: 1944
Score: 3453.00 Matches: 647
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-086-464-2 (1-647) x AAF77094 (1-1944)

Qy	1	MetSerAlaProSerProGlyThrGlySerProProSerProSerAsnSerThr	20
Db	1	ATGTCTCGGCGCGCTCTCCGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA	60
Qy	21	ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProPro	40
Db	61	ACCACCACTCTCTCCAGCTTCGCTCTCCACACACACACCTTCTCTCTCCGCGC	120
Qy	41	ProSerThrThrProThrSerProProSerSerArgSerThrProSerAlaPro	60
Db	121	CCATCCATATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	180
Qy	61	ProSerProThrProSerThrProGlySerProProProLeuProGlnProSerPro	80
Db	181	CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCT	240
Qy	81	ProAlaProThrProGlySerProProAlaProValThrProProThrArgSerPro	100
Db	241	CCCGCTCCAACTACGCGGGATCTCCACCGCACCTGTACTCTCTCTCTCTCTCTCT	300
Qy	101	ProProSerValProGlyProProSerSerProSerArgGluGlyGlySerProArgPro	120
Db	301	CCACCTTCAGTCCAGGACCCCGTCCAAATCTTCACGCGAAGGAGGATCTCTCTCG	360
Qy	121	ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValVal	140
Db	361	CCATCT	420
Qy	141	GlyLeuAlaLeuGlyValAlaLeuLeuValLeuValLeuLeuLeuLeuLeuLeuLeu	160
Db	421	GGAAATCGGCATCGGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	480
Qy	161	LysLysLysArgArgAspGluGluAspAlaTyrValProProProProProProPro	180
Db	481	AAGAAGAACGACGAGGAGACCTTACGGTGACAGCAGCAACATGGCGGCACAAACG	540
Qy	181	GlyProLysAlaGlyProTyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	200
Db	541	GGTCCCAAGCCGAGGAGCTTACGGTGACAGCAGCAACATGGCGGCACAAACG	600
Qy	201	ThrProProSerAspHisValValThrSerLeuProProProProProProProPro	220
Db	601	ACACCACTCATGATCT	660
Qy	221	ProArgGlnProProProProProProProProProPheMetSerSerGlyGlySer	240
Db	661	CCACGGCAACCT	720

Qy	241	TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLys	260
Db	721	TACTCGGACCGTCCAGTTCTCTCTCCACCGTCTCCAGGGCTTGTGTAGGCTTCTC	780
Qy	261	SerThrPheThrThrGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu	280
Db	781	AGCACTTTTCATACGAGGAGCTAGTAGACCAATGGTTCTCCGAGCGCAACTTG	840
Qy	281	LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal	300
Db	841	TTAGACAAAGCGGGTTCGGTTACGTCACAAAGTGTGTTCCTCTAGTGGGAAGAAGT	900
Qy	301	AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal	320
Db	901	GCTGTGAAGCAGTTGAAAGTTGGGAGTGTTCAGGGAGAGAGGAGTTTCAGGAGG	960
Qy	321	GluLeuLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla	340
Db	961	GAGATCATCAGCAGAGTTCCACACAGGCATCTGGTGTCTCTTGTGTGTATTCATC	1020
Qy	341	GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu	360
Db	1021	GGTCCAAAAGATTGCTTGTCTATGAGTTTGTCTTAACAACATCTCGAGCTTCACCT	1080
Qy	361	HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer	380
Db	1081	CATGGGAGGCGGCTTACAAATGGAATGGAGCACAGATTGAAGATTGCTCTTGGATCT	1140
Qy	381	AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle	400
Db	1141	GCTAAAGGACCTTCTTATCTTCAAGAGATTGCAATCCTAAATCAATTCACCGTATAT	1200
Qy	401	LysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLeu	420
Db	1201	AAGCTTCAACACATATTGATAGATTCAAGTTGAAGCTAAGTTGCTGATTTTGTCTT	1260
Qy	421	AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly	440
Db	1261	GCTAAGATTGCTCTGTATACAAACACGCATCTATCAACACGTCGTGATGGAACTTGG	1320
Qy	441	TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer	460
Db	1321	TACTTGGCTCCGGAATACGCTGCAAGCGGAAAGCTCACGGAGAAAGTCTGAGCTTTCT	1380
Qy	461	PheGlyValValLeuLeuGluLeuLeuThrGlyArgArgProValAspAlaAsnVal	480
Db	1381	TTTGGCGTTGTCTTTGGAGCTCATTTACTGAGAGCTGACCCCGTTGATGCCAACATGTC	1440
Qy	481	TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln	500
Db	1441	TATGTAGATGACAGCTTAGTTGACTGGGACGACCATTTGCTTAACCGAGCATCTGAGCA	1500
Qy	501	GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet	520
Db	1501	GGAGACTTTGAGGGTTTAGCTGATGCAAAAGATGAATAATGGGTATGACAGAGAGATG	1560
Qy	521	AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg	540
Db	1561	GCTCGCATGTTGTCTTGTCTGCGGCTGTGTGTGCCATTGAGCTCGCGCAGACCTCGC	1620
Qy	541	MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly	560
Db	1621	ATGAGCCAGATTGTCGTGCGTGTAGAAGAAATGTATCACTGTCTAGATCTTAAACGAAG	1680
Qy	561	MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer	580
Db	1681	ATGAGACGAGTCAAAAGCAATGTATACAGCTCATACGGGAAAGCAGGATTTATGACTCG	1740
Qy	581	SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr	600
Db	1741	AGCCAGTACATGAAGACATGAAGAGTTTAGGAAAATGGCATTGGCACTTCAAGAGTAC	1800
Qy	601	AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer	620

Db 1801 AACGCCAGGGTGAAGTACAGTAATCCACAGTACGACTATGACCGTCTGGTTCA 1860
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysLeuLysArgThrGlyGln 640
Db 1861 AGCAGCAGGGCCAAACACACACGGAATGGAGATGGGAGATTAGAGACCGGTTCAG 1920
Qy 641 GlyTyrSerGlyProSerLeu 647
Db 1921 GGTATTAGTGGACCTTCTCTT 1941

RESULT 2
ACF36548
ID ACF36548 standard; cDNA; 1944 BP.
XX
AC ACF36548;
XX
DT 18-DEC-2003 (first entry)
XX
DE B. napus PERK1 receptor kinase encoding cDNA.
XX
KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ss.
XX
OS Brassica napus:
XX
FH Key Location/Qualifiers
FT CDS 1..1944
FT /tag= a
FT /transl_except= (pos: 1332..1334, aa: Phe)
FT /product= "PERK1"
FT /notes= "Proline-rich Extensin-like Receptor Kinase"
XX
PN WO2003072763-A1.
XX
PD 04-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-CA000274.
XX
PR 28-FEB-2002; 2002CA-02373903.
PR 28-FEB-2002; 2002US-00086464.
XX
PA (GORI/) GORING D.
PA (SILV/) SILVA N.
PA (HAFF/) HAFFANI Y Z.
XX
PI Goring D, Silva N, Haffani YZ;
XX
DR WPI; 2003-712727/67.
DR P-PSDB; ABR82937.
XX
PT Producing a transgenic plant having an increased plant resistance, plant
PT growth or seed production comprises transforming a plant with a nucleic
PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
PT activity.
XX
PS Claim 3; Fig 1A; 123pp; English.
XX
CC The invention relates to producing a transgenic plant having increased
CC plant height, number of branches, number of seed pods and/or seed
CC production compared to a non-transgenic plant, and/or quicker flowering
CC or later senescence compared to a non-transgenic plant. The method
CC involves transforming a plant with a vector including a proline-rich
CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
CC acid molecule having PERK activity. The method, as well as the PERK
CC nucleic acid molecule and polypeptide, are useful in increasing plant
CC resistance to wounding and pathogens and in increasing plant growth and
CC seed production. The nucleic acid molecule and polypeptide may also be
CC used in producing transgenic plants or transgenic host cells. The present
CC sequence represents a cDNA encoding a B. napus PERK1 receptor kinase
CC polypeptide
XX
SQ Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.62e-93 Length: 1944
Score: 3453.00 Matches: 647
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-086-464-2 (1-647) x ACF36548 (1-1944)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProSerAsnSerThr 20
Db 1 ATGTCTCGGCGCGCTCTCCGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA 60
Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProProPro 40
Db 61 ACCACCACCTCTCTCCAGCTTCGGCTCTCTCCACACACACCTTCTTCTCTCCGCG 120
Qy 41 ProSerThrIleProThrSerProProProSerSerArgSerThrProSerAlaProPro 60
Db 121 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Qy 61 ProSerProProThrProSerThrProGlySerProProProLeuProGlnProSerPro 80
Db 181 CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCT 240
Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100
Db 241 CCGCTCCCACTACGCGCGGATCTCCACCGCGACTGTACTCTCTCTCTCTCTCTCTCT 300
Qy 101 ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro 120
Db 301 CCACCTTCAGTCCGAGGACCCAGCGTCCAATCTTACGCGAAGGAGGATCTCTCTCGAC 360
Qy 121 ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140
Db 361 CCATCTCTCTCTCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Qy 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160
Db 421 GGATGCCCATCGAGGAGTGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Qy 161 LysLysLysArgArgAspGluGluAspAlaTyrTyrValProProProProProProPro 180
Db 481 AAGAAGAAACGCGAGAGACGAAGAAGATGCTTACTATGTCTCTCTCTCTCTCTCTCTCT 540
Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnThrArgGlnGlnAsnAla 200
Db 541 GGTCCCAAGCCGAGGAGACCTTACGGTGGACAGCAGCAACAATGGCGGCAACAAACGCA 600
Qy 201 ThrProSerSerAspHisValValThrSerLeuProProProProLysAlaProSerPro 220
Db 601 ACACCCAGTCAGATCATGTCTGTGAGCTACTACCCACACCACTAAGGCTCCATCTCCA 660
Qy 221 ProArgGlnProProProProProProPheMetSerSerSerGlyGlySerAsp 240
Db 661 CCACGGCAACCTCTCTCCACCTCCACCGCTTTCATGAGCAGCAGCGCGGCTCCGAC 720
Qy 241 TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLys 260
Db 721 TACTCGGACCGTCCAGTCTCTCTCCACCGCTCTCCAGGCTGTGTGTGTGTGTGTGTGT 780
Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280
Db 781 AGCATTTACATACGAGGAGCTAGTAGAGCCACCAATGGTTTCTCCGAGGCGAATCTG 840
Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300
Db 841 TTAGGACAAGGCGGTTCTGGTTACGTTCAGTGCAAAAGGTGTGTGTGTGTGTGTGTGTGT 900
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320

```
Db 901 GCTGTGAAGCAGTGTGAAAGTTGGAGTGGTCTCAGGAGAGAGAGGAGCTTTTCAGGCAGAGGTT 960
Qy 321 GluIleIleSerArgValHisHsiArgHisLeuValSerLeuValGlyTyrCysIleIle 340
Db 961 GAGATCATCAGCAGAGTTCACACACAGGATCTGGTGTCTCTTGTGGTTATTGCAATGCC 1020
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu 360
Db 1021 GGTGCAAAAGATTGCTTGTCTATGAGTTGTCTTCAACAACAATCTCAGCTTCACCTC 1080
Qy 361 HisGlyGluArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db 1081 CATGGCAGGAGCGGCTTACAATGGAATGGAGCACAGATTGAAGATTGCTCTTGGATCT 1140
Qy 381 AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 1141 GCTAAAGGACTTCTTATCTTCATGAAGATTGCAATCTTAAATCATTCACCGTGATATC 1200
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 1201 AAGGCTTCACACATATTGATAGATTCAAGTTTGAAGCTAAGTTGCTGATTTGGTCTT 1260
Qy 421 AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 1261 GCTAAGATTGCTTCTGATACAAACACGATGTATCAACACGTTGATGGAAACCTTTGGG 1320
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 1321 TACTTGGCTCCGGAATACGTCACAGCGGAAGCTCACGGAGAAGCTCTGAGCTTTTCTCA 1380
Qy 461 PheGlyValValLeuLeuGluIleThrGlyArgArgProValAspAlaAsnVal 480
Db 1381 TTTGGCGTTGTCTTTGGAGCTCATTAAGTTTGAAGCTAAGTTGCTGATTTGGTCTT 1440
Qy 481 TyrValAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500
Db 1441 TATGTAGATGACAGCTTAGTTGACTGGGCACGACCATTTGCTTAACCGAGCATCTGAGCA 1500
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluMet 520
Db 1501 GGAGACTTTGAGGGTTTACGTGATCAAGATGAATAATGGGTATGACAGAGAGATG 1560
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1561 GCTCGATGGTTGCTTGTCTGCGGCTGTGTTCGCCATTGAGCTCGCGCAGACTCGC 1620
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1621 ATGAGCCAGATTGTGCGTGTAGAGGAAATGTATCACTGTCAATCTTAAACGAAGGG 1680
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlySerThrAspTyrAspSer 580
Db 1681 ATGAGACCAAGGTCAAAGCAATGTATACAGCTCATACGGAGGAAGCACCAGTTATGACTCG 1740
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1741 AGCCAGTACATGAGACATGAAGAAGTTTAGGAAATGGCACTTGGAACTCAAGATAC 1800
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1801 AACGCCACGGGTGAGTACAGTAATCCGACCACTATGGAAGTGGAGTGGGAGATTAAGAGAACCGGTGAG 1860
Qy 621 SerSerGluGlyGlnThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1861 AGCAGCGAGGGCCAAACACACGCGAAATGGAGATGGGGAAGATTAAAGAGAACCGGTGAG 1920
Qy 641 GlyTyrSerGlyProSerLeu 647
Db 1921 GGTATAGTGGACCTTCTT 1941
```

RESULT 3

ACF36549

ID ACF36549 standard; cDNA; 2189 BP.

```
XX ACF36549;
XX 18-DEC-2003 (first entry)
XX B. napus PERK1 receptor kinase encoding cDNA.
XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
XX wound; pathogen resistance; plant growth; seed production; gene; ss.
XX Brassica napus.
XX Key Location/Qualifiers
XX CDS 1..2189
XX /*tag= C
XX /note= "ABR82938; this protein contains the amino acids
XX corresponding to 5' and 3'UTR regions, though only the
XX relevant aa residues (ABR82937) is used in the invention"
XX 5'UTR 1..96
XX /*tag= a
XX CDS 97..2040
XX /*tag= b
XX /transl_except= (pos: 1332..1334, aa: Phe)
XX /product= "PERK1"
XX /note= "ABR82937; proline-rich Extensin-like Receptor
XX Kinase"
XX 3'UTR 2041..2189
XX /*tag= d
XX WO2003072763-A1.
XX 04-SEP-2003.
XX 28-FEB-2003; 2003WO-CA000274.
XX 28-FEB-2002; 2002CA-02373903.
XX 28-FEB-2002; 2002US-00086464.
XX (GORI/) GORING D.
XX (SILV/) SILVA N.
XX (HAFF/) HAFFANI Y Z.
XX Goring D, Silva N, Haffani YZ;
XX WPI: 2003-712727/67.
XX P-PsDB; ABR82937, ABR82938.
XX Producing a transgenic plant having an increased plant resistance, plant
XX growth or seed production comprises transforming a plant with a nucleic
XX acid molecule having a Proline-rich Extensin-like Receptor Kinase
XX activity.
XX Disclosure; Fig 1D; 123pp; English.
XX The invention relates to producing a transgenic plant having increased
XX plant height, number of branches, number of seed pods and/or seed
XX production compared to a non-transgenic plant, and/or quicker flowering
XX or later senescence compared to a non-transgenic plant. The method
XX involves transforming a plant with a vector including a Proline-rich
XX Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
XX acid molecule having PERK activity. The method, as well as the PERK
XX nucleic acid molecule and polypeptide, are useful in increasing plant
XX resistance to wounding and pathogens and in increasing plant growth and
XX seed production. The nucleic acid molecule and polypeptide may also be
XX used in producing transgenic plants or transgenic host cells. The present
XX sequence represents a cDNA encoding a B. napus PERK1 receptor kinase
XX polypeptide
XX Sequence 2189 BP; 544 A; 592 C; 494 G; 559 T; 0 U; 0 Other;
XX Alignment Scores: 5e-93 Length: 2189
XX Pred. No.: 3453.00 Matches: 647
XX Score:
```

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-086-464-2 (1-647) x ACF36549 (1-2189)

[illegible]

XX A. thaliana PERK1 protein encoding genomic DNA.
DE PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ds.
XX Arabidopsis thaliana.
OS
XX WO2003072763-A1.
PN
XX 04-SEP-2003.
PD
XX 28-FEB-2003; 2003WO-CA000274.
PF
XX 28-FEB-2002; 2002CA-02373903.
PR
XX 28-FEB-2002; 2002US-00086464.
PR
XX (GORI/) GORING D.
PA (SILV/) SILVA N.
PA (HAF/) HAFANI Y Z.
XX Goring D, Silva N, Haffani YZ;
XX WPI; 2003-712727/67.
DR
XX Producing a transgenic plant having an increased plant resistance, plant
PT growth or seed production comprises transforming a plant with a nucleic
PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
PT activity.
XX
XX Disclosure; Fig 21d; 123pp; English.
PS
XX The invention relates to producing a transgenic plant having increased
CC plant height, number of branches, number of seed pods and/or seed
CC production compared to a non-transgenic plant, and/or quicker flowering
CC or later senescence compared to a non-transgenic plant. The method
CC involves transforming a plant with a vector including a proline-rich
CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
CC acid molecule having PERK activity. The method, as well as the PERK
CC nucleic acid molecule and polypeptide, are useful in increasing plant
CC resistance to wounding and pathogens and in increasing plant growth and
CC seed production. The nucleic acid molecule and polypeptide may also be
CC used in producing transgenic plants or transgenic host cells. The present
CC sequence represents an A. thaliana PERK1 DNA (At3g24550)
XX
SQ Sequence 1959 BP; 496 A; 514 C; 425 G; 524 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: Length: 1959
Score: 2936.50 Matches: 574
Percent Similarity: 90.61% Conservative: 24
Best Local Similarity: 86.97% Mismatches: 41
Query Match: 85.04% Indels: 21
Dbs: 9 Gaps: 12

US-10-086-464-2 (1-647) x ACF36557 (1-1959)

QY 1 MetSerAlaProSerProGlyThr-----GlySerProProSerProSer 17
Db 1 ATGTCACAGCGCGCTCTCCAGGCACTACTCCATCACCATCTCCACCGTCTCTCCACA 60
QY 18 AnSerThrThrThrThrProProAla-----SerAlaProProThrThrPro 35
Db 61 AACTGCACAACACCACTCTCTCTCCAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 36 SerSerProProPro-----ProSerThrThrThrThrThrThrThrThrThrThr 54
Db 121 TCTTCT 180
QY 55 ThrProSerAlaProProProSerProProThrProSerProSerProProPro 74
Db 181 CCT 228

QY 75 LeuProGlnProSerProAlaProThrThrProGlySerProProAlaProValThr 94
Db CTTCTCAACCTTCCCGCTCCCATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 285
QY 95 ProProThr---ArgAsnProProSerValProGlyProProSerAsnProSerArg 113
Db CCCTCAAAACCTCGAAGCCCTCCATCTCTTAACCAAGGACCACCAAAACCTCCCTCA 342
QY 114 GluGlyGlySerProArgProSerSer-----ProSerProProSerProSer 131
Db ---GGATCTACTCTAGAACTCCATCAACACGCTCGCGCCCTCTGATCTCTCC 399
QY 132 AspGlyLeuSerThrGlyValValValValValValValValValValValValVal 151
Db GATGATTGTCTACCGGAGTTGTGTAGTAATCGCATTTGAGGAGTCTCTCTCTCTCT 459
QY 152 IleValThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
Db ---:---:---:---:---:---:---:---:---:---:---:---:---:---:---: 519
QY 170 AspAlaTyrTyrValProProProProProProProProProProProProPro 189
Db ---:---:---:---:---:---:---:---:---:---:---:---:---:---:---: 579
QY 190 GlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 208
Db GGTCAACAACAGTAGTTGGCAACAACAAACGCGTCACGCGCGTCAGATAATCATGTAG 639
QY 209 ThrSerLeuProProProProProProProProProProProProProProPro 228
Db ---:---:---:---:---:---:---:---:---:---:---:---:---:---:---: 696
QY 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProVal 247
Db CCACCACGAGATTCATGAGTAGCAGTGTGTCTGACTATTCGATCTTCGGTCTCT 756
QY 248 ProProProSerProGlyLeuValValValValValValValValValValVal 267
Db CTCTCCACCATCTCCAGGCTGTGTAGGCTTTCTAAAGACACTTTCATCTATGAGGAG 816
QY 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlyGlyPheGly 287
Db ---:---:---:---:---:---:---:---:---:---:---:---:---:---:---: 876
QY 817 TTGTCGAGAGCTACTAATGGCTTCTCTGAGGCTAATTTGTTAGGACAGGAGGTTTGG 876
QY 288 TyrValHisGlyGlyValValValValValValValValValValValValVal 307
Db ---:---:---:---:---:---:---:---:---:---:---:---:---:---:---: 936
QY 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuLeuSerArg 327
Db GGTAGTGTCTCAGGAG 996
QY 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
Db ---:---:---:---:---:---:---:---:---:---:---:---:---:---:---: 1056
QY 348 TyrGluPheValProAsnAsnLeuLeuLeuHisLeuHisGlyGlyGlyArgProThr 367
Db 1057 TATGATTTGTTCCAAACAACATCTTGAGTTTACCTCCATGTTAGGAGCGGCTACG 1116
QY 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyr 387
Db 1117 ATGGAATGGAGTACTAGATTGAAGATTGCTCTGTGATCTGCTTAAAGAGACTTTCAT 1176
QY 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407
Db 1177 CATGAAGATTGCAATCCGAAATCATTCACCGTGTATTAAGCGCTCAACATATGATT 1236
QY 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAsp 427
Db 1237 GATTTCAAATTTGAAGCTAAGTTGCTGCTCTGCTCTGCTGCTGCTGCTGCTGCT 1296
QY 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447

```
Db      1297 AACACTCATGTATCTACACGGTGATGGGAACCTTTGGGTATTGGCTCCGGAAATGCT 1356
Qy      448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db      1357 GCAGTGGAAAGCTCACAGAAAGTCTGACGTTTTCTCAATTGGCGTTGTACTTTTGAA 1416
Qy      468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db      1417 CTTATTACTGGAGCGCCCTGTTGATGCGCAACAATGCTATGTATGATGATGACAGCTTAGTT 1476
Qy      488 AspTrrAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db      1477 GACTGGCAGCACCATTGCTTAACCGAGCATCTGAGGAAGAGATTITGAGGGTTTGCT 1536
Qy      508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db      1537 GATTCAAAGATGGTAAATGATGATGACAGAGAGAGATGGCTGCATGCTTGGCT 1596
Qy      528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db      1597 GCGCTTGTGTTCGCCATTTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTACGGCG 1656
Qy      548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db      1657 TTAGAGGAATGATCGCTGCTGTCTGATCTTAACGAAGGATGAGACCGGGTCACAGCAAC 1716
Qy      568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db      1717 GTATACAGCTCATATGGAGGAGACAGACTATGACAGAGCCCAATACACGACGACATG 1776
Qy      588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db      1777 ATAAAGTTTAGGAAATGGCTCTTGGAACTCAGATACAGATACGGCACACCGCGAGTACAGT 1836
Qy      608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnThrThr 627
Db      1837 AATCCAAACAGTGACTAGGACTGTACCCGCTCGTTCAAGCAGTGAAGGTCAAGCCACA 1896
Qy      628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
Db      1897 CGAGAAATGAGATGGGAAGATTAAAGAAACCGGTCAAGTTATAGTGGACCTCTCTT 1956

RESULT 5
AAC36968
ID AAC36968 standard; DNA; 2324 BP.
AC AAC36968;
XX
XX 17-OCT-2000 (first entry)
DT
DE
DE
DE
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
```

```
PR 08-APR-1999;
PR 16-APR-1999;
PR 19-APR-1999;
PR 21-APR-1999;
PR 23-APR-1999;
PR 23-APR-1999;
PR 28-APR-1999;
PR 30-APR-1999;
PR 30-APR-1999;
PR 05-MAY-1999;
PR 06-MAY-1999;
PR 06-MAY-1999;
PR 07-MAY-1999;
PR 11-MAY-1999;
PR 14-MAY-1999;
PR 14-MAY-1999;
PR 14-MAY-1999;
PR 18-MAY-1999;
PR 19-MAY-1999;
PR 20-MAY-1999;
PR 21-MAY-1999;
PR 24-MAY-1999;
PR 25-MAY-1999;
PR 27-MAY-1999;
PR 28-MAY-1999;
PR 01-JUN-1999;
PR 03-JUN-1999;
PR 04-JUN-1999;
PR 07-JUN-1999;
PR 08-JUN-1999;
PR 10-JUN-1999;
PR 10-JUN-1999;
PR 14-JUN-1999;
PR 16-JUN-1999;
PR 16-JUN-1999;
PR 17-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 21-JUN-1999;
PR 22-JUN-1999;
PR 23-JUN-1999;
PR 23-JUN-1999;
PR 28-JUN-1999;
PR 28-JUN-1999;
PR 30-JUN-1999;
PR 01-JUL-1999;
PR 01-JUL-1999;
PR 02-JUL-1999;
PR 06-JUL-1999;
PR 08-JUL-1999;
PR 09-JUL-1999;
PR 12-JUL-1999;
PR 13-JUL-1999;
PR 14-JUL-1999;
PR 15-JUL-1999;
PR 16-JUL-1999;
PR 16-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
```

```
99US-0128714P.
99US-0129845P.
99US-0130077P.
99US-0130449P.
99US-0130510P.
99US-0130891P.
99US-0131449P.
99US-0132048P.
99US-0132407P.
99US-0132484P.
99US-0132485P.
99US-0132486P.
99US-0132487P.
99US-0132863P.
99US-0134256P.
99US-0134218P.
99US-0134219P.
99US-0134221P.
99US-0134370P.
99US-0134768P.
99US-0134941P.
99US-0135124P.
99US-0135353P.
99US-0135629P.
99US-0136021P.
99US-0136392P.
99US-0136782P.
99US-0137222P.
99US-0137528P.
99US-0137502P.
99US-0137724P.
99US-0138094P.
99US-0138540P.
99US-0138847P.
99US-0139119P.
99US-0139452P.
99US-0139453P.
99US-0139492P.
99US-0139454P.
99US-0139455P.
99US-0139456P.
99US-0139457P.
99US-0139458P.
99US-0139459P.
99US-0139460P.
99US-0139461P.
99US-0139462P.
99US-0139463P.
99US-0139750P.
99US-0139763P.
99US-0139817P.
99US-0139899P.
99US-0140353P.
99US-0140354P.
99US-0140895P.
99US-0140823P.
99US-0140991P.
99US-0141287P.
99US-0141842P.
99US-0142154P.
99US-0142055P.
99US-0142390P.
99US-0142803P.
99US-0142920P.
99US-0142977P.
99US-0143542P.
99US-0143624P.
99US-0144005P.
99US-0144085P.
99US-0144086P.
99US-0144325P.
99US-0144331P.
99US-01444332P.
```

PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144353P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 21-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145114P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 28-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161922P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161992P.
PR 04-AUG-1999;	99US-0147204P.	PR 28-OCT-1999;	99US-0161993P.
PR 05-AUG-1999;	99US-0147302P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147192P.		
PR 06-AUG-1999;	99US-0147260P.	Alignment Scores:	
PR 06-AUG-1999;	99US-0147303P.	Pred. No.:	1.88e-76
PR 06-AUG-1999;	99US-0147416P.	Score:	2885.00
PR 09-AUG-1999;	99US-0147493P.	Percent Similarity:	90.30%
PR 09-AUG-1999;	99US-0147935P.	Best Local Similarity:	86.36%
PR 10-AUG-1999;	99US-0148171P.	Query Match:	83.55%
PR 11-AUG-1999;	99US-0148313P.	DB:	3
PR 12-AUG-1999;	99US-0148314P.		
PR 12-AUG-1999;	99US-0148341P.	US-10-086-464-2 (1-647) x AAC36968 (1-2324)	
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.	Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProSerProSerProSer 17	
PR 16-AUG-1999;	99US-0149368P.	Db 105 ATGTCCACAGCGCGTCTCCAGGCACCTACTCCATCATCTCCACCGTCTCTCCACCA 164	
PR 17-AUG-1999;	99US-0149175P.	Qy 18 AsnSerThrThrThrProProProAla-----SerAlaProProProThrThrPro 35	
PR 18-AUG-1999;	99US-0149428P.	Db 165 AACTCGACACACCACTCTCTCCAGACGCTTTCTCTCTCTCCACCACTCTCTCT 224	
PR 20-AUG-1999;	99US-0149722P.	Qy 36 SerSerProProPro---ProSerThrThrThrProThrSerProProProSerSerArgSer 54	
PR 20-AUG-1999;	99US-0149723P.	Db 225 TCTTCTCTCTCTCGTCCGCGTCAACTAATCACTCTCTCTCTCTCTCTCTCTCTCTTTA 284	
PR 23-AUG-1999;	99US-0149902P.	Qy 55 ThrProSerAlaProProProSerProProThrProThrProSerThrProGlySerProProPro 74	
PR 23-AUG-1999;	99US-0149930P.	Db 285 CCT 332	
PR 25-AUG-1999;	99US-0150566P.	Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94	
PR 26-AUG-1999;	99US-0150884P.	Db 333 CTTCCTCAACCTTCTCCATCCGCTCCCATCATCTCTCTCTCTCTCTCTCTCTCTCTTCA 389	
PR 27-AUG-1999;	99US-0151065P.	Qy 95 ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113	
PR 27-AUG-1999;	99US-0151066P.	Db 390 CCTCAAACCTCGAAGCCCTCCATCTCTCTCAACGAGGACCACCAACACTCCCTCA--- 446	
PR 30-AUG-1999;	99US-0151080P.	Qy 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131	
PR 31-AUG-1999;	99US-0151303P.	Db 447 ---GGATCTACTCTAGAACCTCCATCAACGCTAAACGCTCGCCACCGCTCTGATCTCTCC 503	
PR 31-AUG-1999;	99US-0151438P.	Qy 132 AspGlyLeuSerThrGlyValvalGlylleAlaIleGlyValAlaLeuVal 151	
PR 01-SEP-1999;	99US-0151930P.	Db 504 GATGGATTGTCTACGGGAGTTGTGTAGGAATCGCCATTGGAGGAGTGCCTATTCTTGT 563	
PR 07-SEP-1999;	99US-0152363P.	Qy 152 IleValThrLeuIleCysLeuLeuCysLysLysLysArgArgArg-----AspGluGlu 169	
PR 13-SEP-1999;	99US-0153070P.	Db 564 ATACGACTCTGATTGTTGTTCTCTCTGTAAGAAGAAACGAGAGAGACGACGATGAA 623	
PR 15-SEP-1999;	99US-0153758P.		
PR 15-SEP-1999;	99US-0154018P.		
PR 16-SEP-1999;	99US-0154033P.		
PR 20-SEP-1999;	99US-0154779P.		
PR 22-SEP-1999;	99US-0155139P.		
PR 23-SEP-1999;	99US-0155486P.		
PR 24-SEP-1999;	99US-0155659P.		
PR 28-SEP-1999;	99US-0156458P.		
PR 29-SEP-1999;	99US-0156596P.		
PR 04-OCT-1999;	99US-0157117P.		
PR 05-OCT-1999;	99US-0157533P.		
PR 06-OCT-1999;	99US-0157865P.		
PR 07-OCT-1999;	99US-0158029P.		
PR 08-OCT-1999;	99US-0158232P.		
PR 12-OCT-1999;	99US-0158369P.		
PR 13-OCT-1999;	99US-0159293P.		
PR 13-OCT-1999;	99US-0159294P.		

QY 170 AspAlaTyrThrValProProProProProGlyProLysAlaGlyGlyProTyrGly 189
 Db 624 GCTGCTTACTATGTT---CTCTCTCCCTCTCTGCTCCCAAGCTGAGGACCTTACGGT 680
 QY 190 GlyGlnGlnGlnTlnPArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208
 Db 681 GGTCAACACACNGATTGTCACACACACACACACCGCTCACGGCGTCAGATATCATGTAGTG 740
 QY 209 ThrSerLeuProProProProProProProProProProProProProProProPro 228
 Db 741 ACGTCATTG---CCACCACCTAAGCCTCCATCTCCACCACGAAACCTCTCCGCGCACCT 797
 QY 229 ProProPro---PheMetSerSerSerGlySerAspTyrSerAspArgProValLeu 247
 Db 798 CCACCACACGATTCATGAGTAGCAGTGGTGTCTGACTATTCGGATCTTCGCGTTCTT 857
 QY 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
 Db 858 CTTCCACATCTCCAGGCTTGTGTAGGCTTTTCAAAAGACATTTCACTTATGAGGAG 917
 QY 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
 Db 918 TTGTCAAGAGCTACTAATAGGCTTCTCTGAGGCTAATTTGTAGGACAAGGAGGTTTGGT 977
 QY 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaLysGlnLeuLysVal 307
 Db 978 TATGTGATAAAGGTATATTGCTAGTGGGAAAGATGCTGTGAAACACAGTTGAAAGCT 1037
 QY 308 GlySerGlyGlnGlyGluArgGlnAlaGluValGluLeuIleIleSerArgValHis 327
 Db 1038 GGGAGTGTTCAGGAGAGAGAGATTTACGGTGAGGTTGAGATCATTAGCAGATTCAT 1097
 QY 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
 Db 1098 CACAGGCAATTGTTCTCTATTGTTGTTATTTGATGCTGATGTTCAAGATGCTTGTGTC 1157
 QY 348 TyrGluPheValProAsnAsnAsnLeuLeuLeuHisLeuHisGlyGluGlyArgProThr 367
 Db 1158 TATGAGTTTGTTCCTAAACAATCTTTGAGTTTCACCTCCATGTAAGGGAAGGCTTACA 1217
 QY 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
 Db 1218 ATGGATGGAGTACTAGATTGAAGATTGCTCTTGATCTGCTAAGGACATTTCAATCTT 1277
 QY 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIle 407
 Db 1278 CATGAAGATTGCAATCCGAAATCATCTCACCGTATATTAAAGCGTCAAAACATATTGATT 1337
 QY 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
 Db 1338 GATTTCAAAATTTGAAGCTAAGGTTGCTGACTTTGCTGTCAGGATTTGCTTCTGATACA 1397
 QY 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
 Db 1398 AACCTCATGTATCTACCGCTGATGGGAACCTTTGGGTATTTGGCTCCAGAAATATGCT 1457
 QY 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
 Db 1458 GCAAGCGGAAGCTCACAGAAAGCTCTGACGCTTTTCTCATTTGTTGTTGTTTTCGAG 1517
 QY 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
 Db 1518 CTTTATAACCGGAGCGCCTGTTGATGCAACAATGCTCTATGTAGATGACAGCTTAGTT 1577
 QY 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyGlyLeuAla 507
 Db 1578 GACTGGGCACACCTTGTCTTAACCGAGCATCTGAGGAAGAGATTTGATGGTTGGCT 1637
 QY 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
 Db 1638 GATTCAAGATGGGTATAGTAGTAGTAG-AGAGAGAGATGCTCCATGGTTGCTTGGCT 1696
 QY 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547

Db 1697 GCAGCTGTGTTCCCATTCAGCTCGCCGAGACCTCGCATGAGCCAGATAGTACGGCG 1756
 QY 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
 Db 1757 TTAGAGGAATATGATCGTGTCTGATCTTAACGAGGATGAGCCGGGTACAGCAAC 1816
 QY 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
 Db 1817 GTATACAGCTCATATGGAGGAAGCACTGACTATGACACAGCAATACAACAGTACATG 1876
 QY 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
 Db 1877 AAGAAGTTTAGGAAATGGCTCTTTGGAACCTCAAGAATACCGCACCAACCGGCGAGTACAGT 1936
 QY 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnGlyGlnThrThr 627
 Db 1937 AATCCAACCACTGACTACGAGCTGTACCCGCTCTGTTCAAGCAGTGAAGGTCAAGCCACA 1996
 QY 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
 Db 1997 CGAGAAATGGAGATGGAAAGATTAAAGAAACCGGTCAAGGTTATAGTGACCCCTCTCTT 2056
 RESULT 6
 ACF36554
 ID ACF36554 standard; DNA; 1812 BP.
 XX AC ACF36554;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE O. sativa PERK protein encoding genomic DNA.
 XX
 KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
 KW wound; pathogen resistance; plant growth; seed production; rice; gene;
 KW ds.
 XX
 OS Oryza sativa.
 XX WO2003072763-A1.
 PN
 PD 04-SEP-2003.
 XX
 PF 28-FEB-2003; 2003WO-CA000274.
 XX
 PR 28-FEB-2002; 2002CA-02373903.
 PR 28-FEB-2002; 2002US-00086464.
 XX
 PA (GORI/) GORING D.
 PA (SILV/) SILVA N.
 PA (HAFF/) HAFFANI Y Z.
 XX
 PI Goring D, Silva N, Haffani YZ;
 XX
 DR WPI; 2003-712727/67.
 XX
 PT Producing a transgenic plant having an increased plant resistance, plant
 PT growth or seed production comprises transforming a plant with a nucleic
 PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
 PT activity.
 XX
 PS Disclosure; Fig 21a; 123pp; English.
 XX
 CC The invention relates to producing a transgenic plant having increased
 CC plant height, number of branches, number of seed pods and/or seed
 CC production compared to a non-transgenic plant, and/or quicker flowering
 CC or later senescence compared to a non-transgenic plant. The method
 CC involves transforming a plant with a vector including a proline-rich
 CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
 CC acid molecule having PERK activity. The method, as well as the PERK
 CC nucleic acid molecule and polypeptide, are useful in increasing plant
 CC resistance to wounding and pathogens and in increasing plant growth and
 CC seed production. The nucleic acid molecule and polypeptide may also be

CC used in producing transgenic plants or transgenic host cells. The present
CC sequence represents an *O. sativa* PERK protein encoding genomic DNA (TIGR
CC Accession No. TC102111)

SQ Sequence 1812 BP; 409 A; 487 C; 512 G; 404 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.6e-45 Length: 1812
Score: 1821.50 Matches: 370
Percent Similarity: 71.62% Conservative: 69
Best Local Similarity: 60.38% Mismatches: 119
Query Match: 52.75% Indels: 55
DB: 9 Gaps: 15

US-10-086-464-2 (1-647) x ACF36554 (1-1812)

```
QY 54 SerThrProSerAlaProProProProProThrProSerThrProGlySerProPro 73
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 TCGTCGCGCTCGTCGCGCGCGCGCTAACCCAGACG-----GCGACACCGCT 51
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 74 ProLeuProGlnProSer-----ProProAlaProThrProGlySerProProAla 91
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 52 CCGGTAACAGACGCGCGCGCACCTCGGGAGACAACACTCGTCGCGCGCGCG 111
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 92 Pro-----ValThrProThrArgAsnProProProSerValProGlyProProSer 109
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 112 CCGGCTCTTTGTCGCGCTCGCTGAGC---CCTCCGCGAGCGCCATCGCCTCGCTG 168
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 110 AsnProSerArgGluGlyGlySerProArgProProSerSerPro-----SerProSer 128
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 169 ACC-----ACGCCCTCTCTCGCGCGCGCGCGCGCGCG 201
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 129 ProSerSerAspGlyLeuSerThrGlyValValGlyValGlyValGlyValGlyVal 148
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 202 GGTCTGAGCTCGGCGCTGACGAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 261
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 149 LeuLeuValLeuValThr-----LeuLeuCysLeuLeuCysLeuLysLysArgArg 166
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 262 GTGCTGCTCTGCTCACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 167 AspGluGluAspAla-----TyrTyr-----Val 174
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 322 CACCACCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 381
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 175 ProProProProProProProProProProProProProProProProProProProPro 194
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 382 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 441
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 195 TrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuProPro 214
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 442 TGGCAGCAAAATGGTCTCCACCTCCACCTGATCATGTGCTG----- 483
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 215 ProLysAlaProSerProProArgGlnProProProProProProProProProProPheMetSer 234
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 484 AAGAAGGTGCTTCGCATCTTCCCTCCACCTCCACCTGCTGCTGCTGCTGCTGCTGCTGCT 543
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 235 SerSerGlyGlySerAspTyrSer-----AspArgProValLeuProProPro 250
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 544 GGTGTTCTGGCTCAAAATTAATCTCCGGTGGCGACACAGCCAACTCTAGTG----- 594
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 251 SerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluLeuAlaArg 270
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 595 TCACCTGGTGTGCTATGGGTTTTCAAGGTGTAATTTTCACTTATGAGAGCTTGTGAGCG 654
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 271 AlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHis 290
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 655 GCGACCGATGGATCTCCGATCTAATCTGCTCGGACAGGCGGGTTTTGGGTATGTTTCCAC 714
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 291 LysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGly 310
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 715 AAAGAGGTCTCCCAATGGGACAGAGTTGTGTGTGAAGCAATGAGAGATGGGATGGG 774
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 311 GlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisArgHis 330
```

```
Db 775 CAGGAGAGCGCTGAATTTACGCGGAGGTTGAGATTATACCGGGTACATCACAAAGAT 834
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 331 LeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPhe 350
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 835 CTTGTAAACATTGGTTGGTTATTCATTTCTGGAGGAGAGAGTGTCTTGTCTATGAGTAT 894
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 351 ValProAsnAsnLeuLeuLeuHisLeuHisGlyGluGlyArgProThrMetGluTrp 370
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 895 GTTCCAAACAACACATTGGAGCTCCACTTGCATGGAGAGCGCCAGCAACCATGGAATGG 954
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 371 SerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLysSerTyrLeuHisGluAsp 390
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 955 CCTACAAGATTACGCAATTCCTTGGGTGCTCGAAGGATTTGGCATATCTTTCATGAAGAT 1014
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 391 CysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLys 410
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1015 TGCATCTCTAAGATCAATTCACCGTGACATAAAGTCAGCAATAATTTCTTTCAGCCGAGA 1074
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 411 PheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHis 430
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1075 TTTGAGGCAAGGTGGCAGATTTCCGACTTGCANAATTGACTTCTGTATAATAACACTCAT 1134
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 431 ValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGly 450
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1135 GTTTCAACACAGAGTAATGGGCACATTTGGGTACCTTGCACAGAGTATGCGTCTTCTGCG 1194
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 451 LysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuLeuThr 470
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1195 CAGCTAACTGAGAAATCAGATGTCTTTCTTCGAGTAATGCTTCTTTCAGCTAATAACT 1254
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 471 GlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuValAspTrpAla 490
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1255 GCGCGCGCTCTGTGAGATCGAACCAATCACAGATGGATGACAGCTTGTGTTGTTGGCA 1314
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 491 ArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyGlyLeuAlaAspAlaLys 510
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1315 AGGCTTTAATGATGCGAGCATCAGATGATGCGCAATTATGATGCTTTAGTGGATCCACGG 1374
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 511 MetAsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAlaAlaCys 530
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1375 CTAGACAGGAGTACATGGTAAACAGATGGCGAGGATGATTGCTTGTGCGAGCTGCTGT 1434
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 531 ValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGly 550
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1435 GTACGCCATTCTGCAGCGCGACGCCACCGATGAGCCAGGTGTTTCGGCGCTTGAAGGC 1494
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 551 AsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSer 570
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1495 GATGTGTCATTGGATGATCTGAACGAGGTGTTCTGCTCTGCCCATAGCCGATTTCTTGGGA 1554
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 571 SerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPhe 590
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1555 TCATAC---AACAGCAACGAGTACGACGCGCCCATTAACATGAGAGCTGAGAGAAATTC 1611
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 591 ArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSerAsnProThr 610
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1612 AGGAAGATGGCATTTGGCAGTGGCAATCTA-----GAGAGCAGCGCAGCAACG 1659
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 611 SerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThrArg----- 628
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1660 CAGCCACCGAGTTCGTCCCGAATCGATCCCGTATCAATGGGGGATGCGCGTCAAAATCG 1719
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 629 -----GluMetGluMetGlyLysIleLysArgThrGly 639
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1720 GAGACGGAGATGGAGATGGGAGTTTGAAGAAAGATGGC 1758
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 7
ACF36555
ID ACF36555 standard; DNA; 1515 BP.
XX
AC ACF36555;
```

DT 18-DEC-2003 (first entry)
 DE Z. mays PERK protein encoding genomic DNA.
 XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
 KW wound; pathogen resistance; plant growth; seed production; maize; gene;
 KW ds.
 XX Zea mays.
 OS WO2003072763-A1.
 PN 04-SEP-2003.
 FD 28-FEB-2003; 2003WO-CA000274.
 XX 28-FEB-2002; 2002CA-02373903.
 PR 28-FEB-2002; 2002US-00086464.
 XX (GORI/) GORING D.
 PA (SILV/) SILVA N.
 PA (HAFF/) HAFFANI Y Z.
 XX Goring D, Silva N, Haffani YZ;
 PI WPI; 2003-712727/67.
 DR
 XX
 PT Producing a transgenic plant having an increased plant resistance, plant
 growth or seed production comprises transforming a plant with a nucleic
 acid molecule having a Proline-rich Extensin-like Receptor Kinase
 activity.
 PT
 XX Disclosure; Fig 21b; 123pp; English.
 PS
 XX The invention relates to producing a transgenic plant having increased
 CC plant height, number of branches, number of seed pods and/or seed
 CC production compared to a non-transgenic plant, and/or quicker flowering
 CC or later senescence compared to a non-transgenic plant. The method
 CC involves transforming a plant with a vector including a proline-rich
 CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
 CC acid molecule having PERK activity. The method, as well as the PERK
 CC nucleic acid molecule and polypeptide, are useful in increasing plant
 CC resistance to wounding and pathogens and in increasing plant growth and
 CC seed production. The nucleic acid molecule and polypeptide may also be
 CC used in producing transgenic plants or transgenic host cells. The present
 CC sequence represents a Z. mays PERK protein encoding genomic DNA (TIGR
 CC Accession No. AY108241)
 XX
 SQ Sequence 1515 BP; 400 A; 330 C; 381 G; 404 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.23e-41 Length: 1515
 Score: 1668.00 Matches: 321
 Percent Similarity: 83.87% Conservative: 43
 Best Local Similarity: 73.96% Mismatches: 59
 Query Match: 48.31% Indels: 12
 DB: 9 Gaps: 4
 US-10-086-464-2 (1-647) x ACF36555 (1-1515)
 QY 221 ProArgGlnProProProProProPheMetSerSerSer-----GlyGly 238
 DB 26 CCTCCACAGCGCCGAGG-CCACCACCACGCGATGCTAAATAGTAGTGGCGGATCTGGT 84
 QY 239 SerAspTyrSerAspArgProValLeuProProSerProGlyLeuValLeuGlyPhe 258
 DB 85 TCTAATTAATCTGGCGGCGAGATCTCTCCACCATCCCTGGAGCTGTCTTGGCTTC 144
 QY 259 SerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAla 278
 DB 145 TCGAAGAGCACATTCACCTACAGAGAGCTGTGTGAGGCTACTGATGATTCGGATGCT 204
 QY 279 AsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLys 298

Db 205 AATCTCTCTGGACAAGGTGGTTTGGCTATGTTTCACAGAGGATTGCTGCCTAATGGCAA 264
 QY 299 GluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAla 318
 Db 265 GAGATTGCTGTAAACAAATTGAAACTGGGAAGTGGCCAGGAGAGCGTGAGTTCCAGGCT 324
 QY 319 GluValGluIleLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCys 338
 Db 325 GAGGTTGAGATTATCAGCCGAGTACATCAACAACACCTTGTGTCTTGGTTGGCTATTGC 384
 QY 339 IleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeu 358
 Db 385 ATTTCTGGAGCGAAGGCTGTGTCTATGAGTTGTCTCCCAATAACACATTGGAATTC 444
 QY 359 HisLeuHisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuValIleAlaLeu 378
 Db 445 CACTTACATCGGAAAGGTCCACCACANTGGAGTGGCTGCTAGATTAAAGATCAGTTG 504
 QY 379 GlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArg 398
 Db 505 GGTGCTGCCAAGGTTTAGCTTATCTTCATGAAGACTGCCATCCAAAGATCATCATCGT 564
 QY 399 AspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPhe 418
 Db 565 GACATAAAGGCATCTAACATTTCTTGGACTTCCAATTTGAAGCTAAGGTGCTGATTTT 624
 QY 419 GlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThr 438
 Db 625 GGACTCCAAAGTTCACCTACTGATAACACACCANTGTATCGACAGAGTAATGGGACC 684
 QY 439 PheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspVal 458
 Db 685 TTTGGGTATTTGGCACCTGAGTATGCATCTCTGGCAAGCTAACAGAAAAATCCGATGA 744
 QY 459 PheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsn 478
 Db 745 TTTTCTCTCGGAGTCATGCTTCTTGAGCTTATTACTGGGCGCGACAGTTGACACAACC 804
 QY 479 AsnValTyrValAspAspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSer 498
 Db 805 CAACATATATGATGATGACACTGGTTGACTGGCAAGGCCATTTACTGTGCGGAGCACTT 864
 QY 499 GluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGlu 518
 Db 865 GAGGATGTCATATATGATGCTTTTAGTGATCCTCGGCTGGGAAAGGACTTCAATCCTAAT 924
 QY 519 GluMetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgArg 538
 Db 925 GAGATGGCAAGAATGATAGCTGTGCAGCTGTCATGTGTAGCCCATTTCTGCACGTCGTCG 984
 QY 539 ProArgMetSerGlnIleValArgAlaLeuGluGlyValSerLeuSerAspLeuAsn 558
 Db 985 CCACGTATGATGATGATGCTGGGCTTTTGGAGGCAATGTGTCTTTGGAGGACCTTAAT 1044
 QY 559 GluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyr 578
 Db 1045 GAAGGTGTTCCGGCTCGCCATAGCCGCTTCTTTGGGTCTATAC---AGCAGCTCCGATTAC 1101
 QY 579 AspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGln 598
 Db 1102 GATTCTGGCCAGTACACACGAGGACATGAAGAAGTTCAAGAAGATGGCATTCACAACAAC 1161
 QY 599 GluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSer 618
 Db 1162 -----TATACCAGCAGCCCAATCAGCGCCCAACAGTGAATATGACAGATACCGTCT 1215
 QY 619 GlySerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThr 638
 Db 1216 GCATCAGCAGCGAGGCGCCACGAGCCAGAGAGATGAGTGGGTGGGTGCAATGAAGAAGGT 1275
 QY 639 Gly-----GlnGlyTyrSerGlyProSer 646

Db 1276 GGCTACAGTGGTGGCTACAGCTCAGGATACAGCGGAGCCTCG 1317

RESULT 8

ACF36556

ID ACF36556 standard; DNA; 1488 BP.

XX AC ACF36556;

XX 18-DEC-2003 (first entry)

XX DE G. max PERK partial EST sequence.

XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;

KW wound; pathogen resistance; plant growth; seed production; soybean; gene;

KW ds.

XX Glycine max.

OS WO2003072763-A1.

PN 04-SEP-2003.

XX 28-FEB-2003; 2003WO-CA000274.

XX 28-FEB-2002; 2002CA-02373903.

PR 28-FEB-2002; 2002US-00086464.

XX (GORI/) GORING D.

PA (SILV/) SILVA N.

PA (HAFF/) HAFFANI Y Z.

XX Goring D, Silva N, Haffani YZ;

PI WPI; 2003-712727/67.

XX Producing a transgenic plant having an increased plant resistance, plant growth or seed production comprises transforming a plant with a nucleic acid molecule having a Proline-rich Extensin-like Receptor Kinase activity.

PT Disclosure; Fig 21c; 123pp; English.

XX The invention relates to producing a transgenic plant having increased plant height, number of branches, number of seed pods and/or seed production compared to a non-transgenic plant, and/or quicker flowering or later senescence compared to a non-transgenic plant. The method involves transforming a plant with a vector including a proline-rich Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic acid molecule having PERK activity. The method, as well as the PERK nucleic acid molecule and polypeptide, are useful in increasing plant resistance to wounding and pathogens and in increasing plant growth and seed production. The nucleic acid molecule and polypeptide may also be used in producing transgenic plants or transgenic host cells. The present sequence represents a G. max PERK EST (expressed sequence tag) partial sequence (TIGR Accession No. TC123698)

XX SQ Sequence 1488 BP; 412 A; 291 C; 348 G; 437 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3,248-40	Length:	1488
Score:	1637.50	Matches:	312
Percent Similarity:	89.67%	Conservative:	44
Best Local Similarity:	78.59%	Mismatches:	36
Query Match:	47.42%	Indels:	5
DB:	9	Gaps:	3

US-10-086-464-2 (1-647) x ACF36556 (1-1488)

QY 253 GlyLeuValLeuGlyPheSerLysSerThrPheThrTyrrGluLeuAlaAraGAlaThr 272

Db 3 GGAAATTCATGGGGTCTCTTAAGAGCATTTCAGTATGAGAGATTGGCAGCGCCAACT 62

QY 273 AsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrrValHisLysGly 292

Db 63 GATGGCTTCTCTGATGCCAACCTCCTTGGACAAGAGGATTGGATATGTGCACAGAGA 122

QY 293 ValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGly 312

Db 123 ATTTCTCCCAACGGCAAGGAGGTGGCAGTGAAGCAATTTGAAGGCTGGGAAGCGGCAAGG 182

QY 313 GluArgGluPheGlnAlaGluValGluLeuSerArgValHisHisArgHisLeuVal 332

Db 183 GAGCGTGAATTCGAAGCTGAAGATTGAGATAATTAGCCGTGCCATCACAAGCACTTGTT 242

QY 333 SerLeuValGlyTyrrCysIleAlaGlyAlaLysArgLeuLeuValTyrrGluPheValPro 352

Db 243 TCTTTGGTTGGATACTGCATCACTGGGTCCAGAGGCTGCTGTTTATGAATTTGTTCCC 302

QY 353 AsnAsnAsnLeuGluLeuHisGlyGlyGlyArgProThrMetGluTrpSerThr 372

Db 303 AACAAACACATTGGAAATTCATTTGCATGGAAGAGGAGACCTACCATGGATTGGCCCCA 362

QY 373 ArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrrLeuHisGluAspCysAsn 392

Db 363 AGACTAGAATTGCTTTAGGATCTGCTAAGGAGACTGGCGTATCTTCATGAAGATTGTCT 422

QY 393 ProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu 412

Db 423 CCTAAGATCATCCATCGTGATATCAAGCTGCCAACATCTCTCGGATTTTAAGTTTGAA 482

QY 413 AlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSer 432

Db 483 GCAAAGGTTGCAGATTTTCGGTCTTGCAAAAGTTTCTTCTGATGCAATACCTCATGTTCT 542

QY 433 ThrArgValMetGlyThrPheGlyTyrrLeuAlaProGluTyrrAlaAlaSerGlyLysLeu 452

Db 543 ACTCGAGTGATGGGACTTTTGGGTATTTGGCTCCAGAAATATGCTTCTAGTGGAAAACTG 602

QY 453 ThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 472

Db 603 ACAGACAAATCAGATGTTTCTCTCCTACGAGTCTATGCTCTCGAGTTAATAACCGGAGA 662

QY 473 ArgProValAspAlaAsnAsnValTyrrValAspSerLeuValAspTrpAlaArgPro 492

Db 663 CGGCGCGTCGATAAAAAATCAAACTTCATGGAGATAGTTTGGTAGACTGGGCTAGGCGCT 722

QY 493 LeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyGlyLeuAlaAspAlaLysMetAsn 512

Db 723 TTGCTCACACGAGCTTTGGAGAGAGATGATTTGATTTATTTGACCAAGGCTCCAG 782

QY 513 AsnGlyTyrrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArg 532

Db 783 AATGACTATGATCTTAATGAGATGGCAGAAATGGTGGCTTCTGCTGCGGCTTGCAATCGT 842

QY 533 HisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552

Db 843 CATTCGGCAAGCGTCGACCAAGGATGAGCCAGGTGTGCCGCGCTCTCGAAGGAGATGC 902

QY 553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrrSerSerTyrr 572

Db 903 TCTCTAGCAGATCTGAATGAAGGAATAGACTGGACACAGCACTATGATAGTTCTCAT 962

QY 573 GlyGlySerThrAspTyrrAspSerSerGlnTyrrAsnGluAspMetLysLysPheArgLys 592

Db 963 ---GAAAGCTCAGATTATGACACTGCACAGTACAAGGAAGACATGAAAAAGTTTCAGGAAA 1019

QY 593 MetAlaLeuGlyThrGlnGluTyrrAsnAlaThrGlyGluTyrrSerAsnProThrSerAsp 612

Db 1020 ATGGCAITGGGAACCTCAGAGGTATGGTCAAGCAGTACAGTACAGTCCGCTACAGTACG 1079

QY 613 TyrGlyLeuTyrrProSerGlySerSerSerGluGly-----GlnThrThrArgGlu 629

Db 1080 TATGGTTTAACCCCATCAGGCTCAAGTAGTCAAGCAGACAGCCGCCCAAAACCACAAGGAA 1139

QY 630 MetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrrSerGlyProSer 646


```
QY 368 etGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuH 388
D 1302 TGGAGTCTCCACTAGTTCGCTATCGCTTAGTTCGCTGCGAAGAGACTCGCTTACCTTC 1361
QY 388 isGluAspCys----- 391
D 1362 ACGAAGACTGTAAGTTTAAACATTCAACCATCTCATTTTCTTAACCAAGTTGCATAAACA 1421
QY 392 -----Asn 392
D 1422 AGAGAAAGCTCTCTCTGACTAGTGTATCTTTTGGCTGAGAAAATGGTGCGAGCCAT 1481
QY 393 ProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu 412
D 1482 CCTCGATCAATCACCGGACATCAAGCTGCAATATCTCTGAGACTTCACTTTCAT 1541
QY 413 Ala----- 413
D 1542 GCTATGGTGATAAAGTCTAGTAGTCTGATTCATCTACGGTTTTTGTGTAAGACTACATTGA 1601
QY 414 -----LysValAlaAspPheGlyLeuAlaLysIleAlaSerAs 426
D 1602 TGACATTTTGCATTTGTTTATTCAGGTGCTGATTTGGATTAGCTAAGTTAAACATCTGA 1661
QY 426 pThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
D 1662 TAAACACACTCATGTATCTACTCTGTGTGTGGAACTTTCGGTAAGCGTTTACCGTAT 1721
QY 440 ----- 440
D 1722 GATAAGATTGTTGCTGACACTCAAGAAACATAAACCCTTTGTAGACTAATCTACTTGGCTTC 1781
QY 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrG 454
D 1782 TTTCCACAACATGTTAGATATCTAGCTCCAGATATGCTTCAAGCGGTAAATTAACCG 1841
QY 454 LuLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgP 474
D 1842 AGAATCCGATGTTTCTCTTACGGAGTTATGTTATGGAACTTATTAATCGAATAACGAC 1901
QY 474 roValAspAlaAsnValThrValAspSerLeuValAspTrp----- 489
D 1902 CGGTTGAT---AATAGCATCACCATGGACGACACCTTAGTAGATTGGGTATTCATGCATG 1958
QY 489 ----- 489
D 1959 TAAATATGATCGGTATATATGTTTTTCGGCTTTTTCGGCTACTAATGATCATGAATA 2018
QY 490 ---AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAsp 508
D 2019 CAGGCTCGGCTCTTATGCTCGCGGCTAGAGATGGAACTTTAATGAGCTCGCAGAT 2078
QY 509 AlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAla 528
D 2079 GCGAGCTTGAAGGCAACTACAAACCGCAAGAAATGGCTCGAATGGTACTTGTGCCGT 2138
QY 529 AlaCysValArgHisSerAlaArgArgProArgMetSer----- 542
D 2139 GCTAGCATTCGATTCGCGGGCTAAACGTCAAAAGATGAG-CCAGGTGAATCAAAATTA 2197
QY 543 -----GlnIleVal 545
D 2198 TAACTAAAAGTCTATTTTGTTCAGAGATAACAAACAAATGTTGTGTTTTCAGATAGTA 2257
QY 546 ArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGln 565
D 2258 AGAGCGTTAAGAGGAGAAAGTGTCTTAGATGCTTTAAACGAAGGTGTGAAGCCAGGACAC 2317
QY 566 SerAsnValThrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGlu 585
D 2318 AGTAAGTTCGCGGTTCATTTGGGAGCAAGCTCGGATTTATAGTCAGACATCTTACATGCA 2377
```

```
QY 586 AspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr-----AsnAla 602
D 2378 GACATGAAGAAATTCAGACAGATAGCTTTTTCGAGCCAAAGAAATTCGCCAGTCAGTGACTGT 2437
QY 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGly 614
D 2438 GAAGGAACATCTAGTATGATTTCTAGAGATATGGGA 2473
RESULT 10
ACF36551
ID ACF36551 standard; DNA; 2520 BP.
XX ACF36551;
XX 18-DEC-2003 (first entry)
XX Arabidopsis PERK1 receptor related protein encoding DNA.
DE PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ds.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX WO2003072763-A1.
PN 04-SEP-2003.
PD 28-FEB-2003; 2003WO-CA000274.
PF 28-FEB-2002; 2002CA-02373903.
PR 28-FEB-2002; 2002US-00086464.
XX (GORI/) GORING D.
PA (SILV/) SILVA N.
PA (HAFF/) HAFFANI Y Z.
XX Goring D, Silva N, Haffani YZ;
XX WPI; 2003-712727/67.
DR P-PSDB; ABR82940.
XX Producing a transgenic plant having an increased plant resistance, plant growth or seed production comprises transforming a plant with a nucleic acid molecule having a Proline-rich Extensin-like Receptor Kinase activity.
XX Disclosure; Fig 12; 123pp; English.
CC The invention relates to producing a transgenic plant having increased plant height, number of branches, number of seed pods and/or seed production compared to a non-transgenic plant, and/or quicker flowering or later senescence compared to a non-transgenic plant. The method involves transforming a plant with a vector including a Proline-rich Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic acid molecule having PERK activity. The method, as well as the PERK nucleic acid molecule and polypeptide, are useful in increasing plant resistance to wounding and pathogens and in increasing plant growth and seed production. The nucleic acid molecule and polypeptide may also be used in producing transgenic plants or transgenic host cells. The present sequence represents a PERK1 polypeptide related protein encoding genomic DNA from A. thaliana (Accession NO. AAD15491)
XX SQ Sequence 2520 BP; 717 A; 552 C; 538 G; 713 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.58e-36 Length: 2520
Score: 1504.00 Matches: 367
Percent Similarity: 53.47% Conservative: 80
Best Local Similarity: 43.90% Mismatches: 146
Query Match: 9 Indels: 244
DB: 9 Gaps: 22
US-10-086-464-2 (1-647) x ACF36551 (1-2520)
```

Qy 3 SerAlaProSerProGlyThrGlySerProProSerProProSerProSerAsnSerThrThr 22
Db TCTGCTCTCCACAAACTCCACCTCTCTCCATCTCCACCGTCTAATAACCAATTCAACC 95
Qy 23 ThrProProAlaSerAlaProProProProProProProProProProProProPro 42
Db ACCCTCTCTCCG-----CCGGCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAAGGA 146
Qy 43 ThrIleProThrSerProProProSerSerArgSerThrProSerAlaProProProSer 62
Db GACTCATCATCATGCCACCTCTCTGATTCCACATCTCCACAGCTCCACAAAGCTCCTTAAC 206
Qy 63 ProProThrProProGlySerProProProProProProProProProProProAla 82
Db CCTCCTTAAT---TCTCTTAATAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 263
Qy 83 ProThrThrProGlySerProProAlaProValThrProProThrArgAsnProProPro 102
Db AGAGAAATGGAGGAACAATGGTGGCAATGATATCTCCACCGTCACCGCGCTCTCTCTCTCT 323
Qy 103 SerValProGlyProProSerAsnProSerArgGluGlyGlySerProArg-----Pro 120
Db TCT-----CCTCCTCTTAGGATTAATGGAGATAATGGTGTAGCAGATCATCGCCA 374
Qy 121 ProSer-----SerProSerProProSerPro-----SerSerAsp 132
Db CCAGGAGACACTGGAGGCTCTCGCTCAGACACCTCTCTCTTAGCGGAGGAAGCAGTGGA 434
Qy 133 Gly-----LeuSerThrGlyValValValValGlyIleAlaIleGly 145
Db GGAGGTGGAGTGAAGAAGTAATACGAATACAGCATCATAGTTGGTGTATAGTCGGA 494
Qy 146 GlyValAlaLeuLeuValIleValLeuIleCysLeuLeuCysLysLysLysArgArg 165
Db GCTGGACTTTTGATGATCGTTCTATTATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 554
Qy 166 ArgAspGluGluAspAlaTyThrValProProProProProProProProGlyProLysAla--- 184
Db -----GACTCCTCTTAC-----CCTGAACCCATGAAGGTGAATAAACAATATAC 596
Qy 185 -----GlyGlyPro 187
Db ACATCTTTATGTTTCAACAATAAGAGCTTAGATTCTTTTCATATAAATTTTCAGGAACCA 656
Qy 188 TyrGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisVal 207
Db ATATCAATACTATGGAACAACAACAACAATGCTTTCACAGAAATTTATCCGAA----- 710
Qy 208 ValThrSerLeuProProProProProLysAlaProSerProPro----- 221
Db TTTGGCACCTTAATTCACAAGGCCAAACCAACAATCTACTGTGTGTGTGGG 761
Qy 222 ArgGln-ProProProProProProProPhenMetSerSerSerGly-----GlySe 239
Db AGCGGTGGACCATCAGCGCTCTCTCTCGCGATGCTTCAAGCGGAGAAATTCCTC 821
Qy 239 rAspTySer-----AspArgProValLeuProProProProGlyLeuValLeuG1 257
Db CATGTACTCAGGCCCATCAGCCCGAGTTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 881
Qy 257 yPheSerLysSerThrPheThrThrGluGluLeuAlaArgAlaThrAsnGlyPheSerG1 277
Db ATTCAACAAGACACTTTACTTACCAAGAGCTTGGCGCTGCAACAGAGGGTTTACCGA 941
Qy 277 uAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyThrValHisLysGlyValLeuProSerG1 297
Db TGCTTAACCTTTTGGACAGGAGGATTTGGGTATGTCTCATTAAGAGGTCTTGGCTAGCGG 1001
Qy 297 YLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheG1 317
Db GAAAGAGTAGCAGTTAAGAGTTTAAAGCGGGTAGCGGACNAGAGAGAGAGGAGTTTCA 1061

Qy 317 nAlaGluValGluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTy 337
Db AGCTGAGGTGCATATCATTAGCCGTGTGCATCATCGTATCTTGTCTTCTTGTGGTGATA 1121
Qy 337 rCysIleAlaGlyAlaLysArgLeuLeuValTyGluPheValProAsnAsnLeuG1 357
Db TTGCATAGCTGATGGACAGAGGATGTGGTTATGAGTTGTCTTCTTAACAAACTTTGGA 1181
Qy 357 uLeuHisLeuHisGly----- 362
Db ATATCATCTTCATGTTAGACCACTTAAACCTTTGAGTACTAAGTTTATTTTCTCTAAT 1241
Qy 363 -----Glu-GlyArgProThrM 368
Db CTATATATTCAAGAAAGTTGTAACTTAATTTGTGTGTAGGGAATAATCTTCCGGTAA 1301
Qy 368 etGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyLeuH 388
Db TGGAGTTCTCCACTAGTTGCGTATGCGCTTAGTGTCTGCGAAGGACTCGCTTACCTTC 1361
Qy 388 iGluAspCys----- 391
Db ACGAAGACTGTAAGTTTAAACATTCACCATCTCTCATTTTCTTAACCAAGTTGCATAAAC 1421
Qy 392 -----Asn 392
Db AGAAAGCT 1481
Qy 393 ProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu 412
Db CTTGGATCATTTACCGGACATCAAGTCTGCAATATCTCTTGGACTTCAACTTTGAT 1541
Qy 413 Ala----- 413
Db GCTATGCTGATAAACTAGTAGCTTGCATTCATCTACGGTTTTTGTGTAAAGACTACATTTGA 1601
Qy 414 -----LysValAlaAspPheGlyLeuAlaLysIleAlaSerAs 426
Db TGACATTTTGCATTTGTATTATTCAGGTGGCTGATTTTGGATTAGCTAAGTTAAACATCTGA 1661
Qy 426 pThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
Db TAACAACATCATGTATCTACTCTGTGTATGGAACTTTTCGGGTAAAGCGTTTACCGTAT 1721
Qy 440 ----- 440
Db GATAAGATTGTCGTGACACTCAAGAAACATAACCTTTTGTAGACTAATCTACTTTTGTCTTC 1781
Qy 441 -----TyrLeuAlaProGluTyAlaAlaSerGlyLysLeuThrG 454
Db TTTTCCAAACATGTGTAGATATCTAGCTCCAGAAATATGCTTCAAGCGGTAAATTTAACCG 1841
Qy 454 lLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgP 474
Db AGAAATCCGATGTTTCTTCTAGCGAGTTATGTATTGGAACCTTATAACTTGAAGAAACGAC 1901
Qy 474 roValAspAlaAsnAsnValTyTrpAspAspSerLeuValAspTrp----- 489
Db CGGTTCAT---ANTAGCATCCATCGACACACCTTAGTAGATTGGGTATTTCATGTCATG 1958
Qy 489 ----- 489
Db TAACATATGATGCTGTATATATATGTTTTCGCCCTTTTTCGCCGTACTAATGATCATGTAATA 2018
Qy 490 ---AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAsp 508
Db CAGGCTCGGCTCTTATGGCTCGCGCTAGAGATGGAACCTTATATGAGCTCGCAGAT 2078
Qy 509 AlaLysMetAsnAsnGlyTyTrpAspArgGluGluMetAlaArgMetValAlaCysAlaAla 528
Db GCGAGGCTTGAAGGCAACTACAAACCCGCAAGAAATGGCTCGAATGGTGAATTTGTGCGCGCT 2138
Qy 529 AlaCysValArgHisSerAlaArgArgProArgMetSer----- 542

Db 2139 GCTAGCATTCGTCATTCGGGGCGTAACCGTCCAAAGATGAG-CCAGGTGAATCAAAATTA 2197
 Qy 543 -----GlnIleVal 545
 Db 2198 TAACTAAAAGTCTATTTTGTGCAGAGAATAACAACAAATGTTGTGGTTTTTCAGATAGTA 2257
 Qy 546 AtgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGln 565
 Db 2258 AGAGGGTTAGAGGAGAGAGTGTCTTAGATGTTTAAAGAGGTGTGAAGCCAGGACAC 2317
 Qy 566 SerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGlu 585
 Db 2318 AGTAAACGTTTACGGGTCAATGGGAGCAAGCTCGGATTATAGTCAGACATCTTACAATGCA 2377
 Qy 586 AspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr-----AsnAla 602
 Db 2378 GACATGAAGAAATTCAGACAGATAGCTTTGTGCGCCCAAGAAATCCCAAGTCAGTGACTGT 2437
 Qy 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGly 614
 Db 2438 GNAGGACATCTAGTAAATGATTCTAGAGATATGGGA 2473

RESULT 11
 AAF77097
 ID AAF77097 standard; DNA; 2820 BP.
 XX
 AC AAF77097;
 XX
 DT 17-MAY-2001 (first entry)
 XX
 DE Arabidopsis gene #3.
 XX
 KW Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200114563-A1.
 XX
 PD 01-MAR-2001.
 XX
 PF 18-AUG-2000; 2000WO-CA000966.
 XX
 PR 19-AUG-1999; 99US-0149466P.
 PR 13-OCT-1999; 99US-0159122P.
 XX
 PA (GORI//) GORING D.
 PA (SILV//) SILVA N.
 XX
 PI Goring D, Silva N;
 XX
 DR WPI; 2001-244305/25.
 XX

PT New proline-rich, extensin-like receptor kinase nucleic acids and
 PT polypeptides useful for increasing plant wounding or pathogen resistance,
 PT or for producing transgenic plants with increased wounding or pathogen
 PT resistance.
 XX
 PS Example; Fig 13; 91pp; English.

CC The present invention relates to proline-rich extensin-like receptor
 CC kinase (PERK). The PERK nucleic acids and polypeptides are useful for
 CC increasing the resistance of plants to wounding and pathogens. These are
 CC also useful for producing transgenic plants with increased wounding and
 CC pathogen resistance compared with a wild type plant, as well as in assays
 CC for identifying and developing compounds to inhibit and/or enhance
 CC polypeptide function directly

SQ Sequence 2820 BP; 819 A; 629 C; 577 G; 795 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2e-35 Length: 2820
 Score: 1479.50 Matches: 367

Percent Similarity: 50.00% Conservative: 99
 Best Local Similarity: 39.38% Mismatches: 151
 Query Match: 42.85% Indels: 318
 DB: 4 Gaps: 28
 US-10-086-464-2 (1-647) x AAF77097 (1-2820)
 Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20
 Db 52 GTGATTCATCTCTGCTCCCTGAAACC---TCAAATGGGACACACCGTCAACGGAACA 108
 Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProProPro 40
 Db 109 TCG-----CCGTCTAATGAGTCATGCGCGCAACACACCTTCTTCCACCACCA 159
 Qy 41 ProSerThrIleProThrSerProProSerSerSerSerThrProSerAlaProPro 60
 Db 160 TCATCA-----ATATCTGCTCTCCCGCAGATATCTCCGTTCTTTTTCACCGCG 210
 Qy 61 ProSerProProThr-----ProSerThrProGlySerProProProLeu 75
 Db 211 CCTGCACCACCAAGCAAGAACGTCACCTCTACATCTCCGTCTCATCGCGCTGT 270
 Qy 76 -----ProGln-----ProSerProProAlaPro-----Thr 84
 Db 271 GTAGCTAATCGTCACCGCAGACTCCAGAGAATCTTCTCCACCTGCACCTGAAGGCTCA 330
 Qy 85 ThrProGlySerProProAla---ProValThrPro-----ProThrArgAsn 99
 Db 331 ACTCTGTACGCCACCTGCACCAACACCGTCGAACCAATACCCGGAAGACCA 390
 Qy 100 ProProPro-----ProGln-----ProSerProProAlaPro----- 102
 Db 391 ACTCTCTCTCTCTGTCGCAATGATGACCGCAACAGAACCAATGGCGGAACACCAAC 450
 Qy 103 -----SerValProGlyProProSerAsnProSerArg-----GluGlyGly 116
 Db 451 AGAGACGGCTCCACACCATCACCGCTCGTCAGGAACAGAACTTCGGTGAAGGCTGC 510
 Qy 117 SerProArgProProSerSerProProSerProProSerSerSerSerGlyLeuSerThr 136
 Db 511 TCACCTTCACCACCTCGGTGATTAAGCCCTCCTCAGAAATAGTGAGATTCAGACTCA 570
 Qy 137 -----GlyValValValGlyIleAlaIleGlyGlyVal 147
 Db 571 TCGGTAATCATCCACAAGCCAACTTGGATTGATTATGGAGTCTTGTAGGACGAGG 630
 Qy 148 AlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysArgArgAsp 167
 Db 631 CTTTGTCTTCTACTTGCACTGTATTGTCATCTGTTGCAACAGGAGAGAGAAGAA 690
 Qy 168 GluGluAspAla-----TyrTyrValProProProProProProGlyProLysAla 184
 Db 691 TCTCTCAGGTCAACCATCATGCACTACTACANTAACAACTCTTATGGAGGACCCCTCA 750
 Qy 185 Gly----- 185
 Db 751 GGTAAATTACAGTTTAGTATACTGGAATTTAATTTGTAGCCTAATGGTGTGTTAGGT 810
 Qy 186 -----GlyProTyrGlyGlyGlnGlnGln 194
 Db 811 TTCAGAACGATCATAGTCTAATGGTTTCTGTAGCTCCATATGGCAAAAGATTAGATT- 869
 Qy 195 TrpArgGlnGlnAsnAlaThr-----ProPro 203
 Db 870 TATAGCTAAAGAGAGAGTTTCATAGTGTAGGTAAATGGTGTGTTATTACAAGGGAACCT 929
 Qy 204 SerAspHisValValThrSerLeuProProProProProProProProProProPro 223
 Db 930 CAAGATCATGTGGTG---AATATGGCTGGTCAAGGAGGTGGGAATTTGGGTCACAGCAA 986
 Qy 224 ProProProProProProProProPheMetSerSerSerGlyGlySerAspTyrSerAsp 243

Db 987 CCTGTCTGTCCT-----CACAGTGATGCTTCCAACTTAACCGGT 1028
Qy 244 ArgProValLeuProProSerProGlyLeuValLeuGlyPheSerLysThrPhe 263
Db 1029 CGAACTGTATACCGTCACTCAA---GCTCAACTCTTGGTCAACAAAGCACATTC 1085
Qy 264 ThrTyrGluLeuLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuGlyGln 283
Db 1086 ACATACGATGAACGTGCATTGCAACAGAAAGTTTCGCTCAGTCAAAATTTCTAGGACAA 1145
Qy 284 GlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLys 303
Db 1146 GGAGATTGGGTATGTTCAATAGGAGTTCGCTAGTGCAAGAAAGTTGCGTGAAG 1205
Qy 304 GlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeile 323
Db 1206 AGTCTTAACTTGAAGTGGCAAGGGGACCGAGTTTCAAGCAGAGGTTGATATCAT 1265
Qy 324 SerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343
Db 1266 AGCCGTGTCCATCATCGTCATCTCGTTCTCTGTGGATATTGCATCTCTGGTGGTCAA 1325
Qy 344 ArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuGluHisLeuHis----- 361
Db 1326 AGACTTTTGGTTATGAGTTTATACCTAAACAACACTCTTGAATTTTCATCTTCATGTTACA 1385
Qy 361 ----- 361
Db 1386 TTCATCTAACAGAAATGTTTCTGTATTAAACAAACCTTTTAAGTATGTTTCTCTTTAAT 1445
Qy 362 -----GlyGluGlyArgProThrMetGluTrpSerThrArgLe 374
Db 1446 CAGGAACATGATGAAATTTTCAGAAAGGGTCTCGCGTTTGGATTGGCCCTACAAGAT 1505
Qy 374 uLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys----- 391
Db 1506 GAAATGTCATGGGATCAGCTAGAGGCTTGATATTTGCATGAGACTGTAGAAAAT 1565
Qy 391 ----- 391
Db 1566 CTTTATCTCACATATTTGCATCAGTTTCTATCTCGCTCTCTACAATATTTGAAAGATTGT 1625
Qy 392 -----AsnProLysIleIleHisArgAspIleLysAlaSerAs 404
Db 1626 ATATGTTTACATCAATATPAGTGCACCCCTCGCATTTATCCACAGAGATATCAAAAGTGCAAA 1685
Qy 404 nileuileAspPheLysPheGluAla----- 413
Db 1686 CATCTCTTGATTTTCAGTTTGTAGAC-CAAGGTATGTGTATATATATCGACTCTGTAC 1744
Qy 413 ----- 413
Db 1745 TACTTTTACTTTCATGTCTCATTTTTTGTGTTTCCATCTCTGTGTCGATGTGTATCAGT 1804
Qy 414 -----LysValAlaAspPheGlyLeuAlaLysIleAlaSerAspTh 427
Db 1805 CTTATTGTGTAATAATATGCGAGTGGCAGATTTGGATTGGCTAAGCTATCTCAAGACAA 1864
Qy 427 rAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
Db 1865 CTATACTCATGTCTCCACTCCGTCATGGGAACCTTTTGGGTAAGCAGCTTTGTAATAATGT 1924
Qy 440 ----- 440
Db 1925 CTCAACTCATCCACACTTATTTAGTTTCTTTTCACTTGTGTTTTTAACATTTTCTTGATCA 1984
Qy 441 --TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheS 460
Db 1985 GATACTTAGCTCCAGAGTATGCATCAGCGGAAAGTTATCCGACAAATCTCATGTGTTTCT 2044
Qy 460 erPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAenV 480
Db 2045 CATTTGGAGTAATGCTTCTTGAGCTCATAAACCGGAAGACCTCTCTGTGATCTAACTGGA- 2103

Qy 480 alTyrValAspAspSerLeuValAspTrp----- 489
Db 2104 --GAAATGGAAGATAGCTTGGTAGATTGAGTTCGGTTCGGCTCTTCGGTTTACTT 2161
Qy 489 ----- 489
Db 2162 GTTTAATCCCAAAACACTTTTCCAAAGCAAAACAGAAACAAATCTTACTATTGTTGTTC 2221
Qy 490 --AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspA 509
Db 2222 AGGCAAGCCCTTGTGTTTGAAGACACTCAAGATGAGATTACACCAATTTGCTGATC 2281
Qy 509 laLysMetAsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAlaAlaA 529
Db 2282 CAGCTCTAGAGCTAAACTACAGTCATCAAGAGATGTTCAAAATGGCTTCTTGTGCAGTG 2341
Qy 529 laCysValArgHisSerAlaArgArgProArgMetSerGlnIle----- 544
Db 2342 CAGCAATCAGCATTTCAAGCAAGAACGCGCTAAGATGAGCCAGGT-TCAAAAAACTCATA 2400
Qy 544 ----- 544
Db 2401 CCACTTGTGTTCTATTGTTATATTTTACTCACAAATTAATCTTGATGATAAATGTGAC 2460
Qy 545 -----ValArgAlaLeuGluG 550
Db 2461 ATACTAATGAATCTTGAACCATGTGTATGTTAAATGAAAAAGATTGTACGAGCCTAGAA 2520
Qy 550 lyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrS 570
Db 2521 GAGATGTCAATGGATGATCTAAGTGAGGGAACAAGACAGGCAACAGCAGTACTTGA 2580
Qy 570 erSerTyrGlySerThrAspTyrAspSerGlnTyrAsnGluAspMetLysLysP 590
Db 2581 GCCCGGGAGGCTGAGCTCAGAGTATGACGCAAGCTCGTACACGGCAGACATGAAAAAGT 2640
Qy 590 heArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSerAsnProt 610
Db 2641 TCAAGAAACTGGCGTTAGAGATAAAGATATCAAGCAGT--GAATATGTTGGA--A 2694
Qy 610 hrSerAspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluM 630
Db 2695 CAAGTGTAGTATGGCTTAAACCTTCTGCTTCAAGTAGTGAA-----GAAA 2739
Qy 630 etGluMetGlyLysIleLysArgThrGlyGln 640
Db 2740 TGAATAGAGCTCAATGAAGCAATCCTCAG 2771
RESULT 12
ID ACF36552
ACF36552 standard; DNA; 2820 BP.
XX ACF36552;
AC ACF36552;
XX
DT 18-DEC-2003 (first entry)
XX
DE Arabidopsis PERK1 receptor related protein encoding DNA.
XX
KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
XX wound; pathogen resistance; plant growth; seed production; gene; ds.
XX
OS Arabidopsis thaliana.
XX
XX
PN WO2003072763-A1.
XX
PD 04-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-CA000274.
XX
PR 28-FEB-2002; 2002CA-02373903.
XX
PR 28-FEB-2002; 2002US-00086464.

PA (GORI/) GORING D.
PA (SILV/) SILVA N.
PA (HAFF/) HAFFANI Y Z.
XX
PI Goring D, Silva N, Haifani YZ;
XX
DR WPI; 2003-712727/67.
DR P-PSDB; ABR82941.
XX
PT Producing a transgenic plant having an increased plant resistance, plant
PT growth or seed production comprises transforming a plant with a nucleic
PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
PT activity.
XX
PS Disclosure; Fig 13; 123pp; English.
XX
CC The invention relates to producing a transgenic plant having increased
CC plant height, number of branches, number of seed pods and/or seed
CC production compared to a non-transgenic plant, and/or quicker flowering
CC or later senescence compared to a non-transgenic plant. The method
CC involves transforming a plant with a vector including a proline-rich
CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
CC acid molecule having PERK activity. The method, as well as the PERK
CC nucleic acid molecule and polypeptide, are useful in increasing plant
CC resistance to wounding and pathogens and in increasing plant growth and
CC seed production. The nucleic acid molecule and polypeptide may also be
CC used in producing transgenic plants or transgenic host cells. The present
CC sequence represents a PERK1 polypeptide related protein encoding genomic
CC DNA from A. thaliana (Accession NO. CNA18823)
XX
SQ Sequence 2820 BP; 819 A; 630 C; 576 G; 795 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,42e-35 Length: 2820
Score: 1471.50 Matches: 366
Percent Similarity: 49.8% Conservative: 99
Best Local Similarity: 39.27% Mismatches: 152
Query Match: 42.62% Indels: 318
DB: 9 Gaps: 28

US-10-086-464-2 (1-647) x ACP36552 (1-2820)

QY 1 MetSerAlaProSerProGlyThrGlySerProProSerProSerProSerThr 20
DB 52 GTGGATTTCATCTCTCCCTCCCTGAAACC---TCAAAATGGGACACCCGCTCAAGGAACA 108
QY 21 ThrThrProProProAlaSerAlaProProProThrThrProSerSerProPro 40
DB 109 TCG-----CCGTCTAATAGTATCGCGCGCAACACCACTTCTTCACCACCA 159
QY 41 ProSerThrIleProThrSerProProProSerSerArgSerThrProSerAlaPro 60
DB 160 TCATCA-----ATATCTGCTCTCTCGCCAGATATCTCGCTCTTTTCACCGCG 210
QY 61 ProSerProProThr-----ProSerThrProGlySerProProProLeu 75
DB 211 CTGCACCAACCAACGAAGAAACGTACACCTCTACATCTCGCTCTCATCGCGCGCTGT 270
QY 76 -----ProGln-----ProSerProAlaPro-----Thr 84
DB 271 GTAGCTAAATCCGTACCGCAGACTCCAGAGATCTCTTCTCCACCTGCACCTGAAGGCTCA 330
QY 85 ThrProGlySerProProAla---ProValThrPro-----ProThrArgAsn 99
DB 331 ACTCTGTAAACGCCACTGCACCAACCAACCGTCGAACCAATCACCAGAAAGACCA 390
QY 100 ProProPro-----Pro-----Pro-----Pro----- 102
DB 391 ACT 450
QY 103 -----SerValProGlyProProProSerProSerProSerArg-----GluGly 116
DB 451 AGAGACGGCTCCACCACTCATCACCGCTCGTCAGGGAACAGAACTTCCGGTGCACGGTGC 510

QY 117 SerProArgProProSerProSerProProSerProSerProSerProSerThr 136
DB 511 TCACCTTCACCACTCGGTGCGAATGAGCCCTCTCAGAAATAGTGAGATTTCAGACTCATCA 570
QY 137 -----GlyValValValGlyIleAlaIleGlyGlyVal 147
DB 571 TCGGTAATCATCCACCAAGCAATTTGAGTTGATTTAGTTCCTTGTAGGAGCAGGG 630
QY 148 AlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysArgArgAsp 167
DB 631 CTTTGTCTTCTACCTGTGAGTGTATTTGCAATCTGTTGCAACAGGAAGAAAGAA 690
QY 168 GluGluAspAla-----TyrTyrValProProProProProGlyProLysAla 184
DB 691 TCTCTCAGGTCAACCATGCCTACTACATAACAATCTTATGGAGGACCCCTCA 750
QY 185 Gly----- 185
DB 751 GGTAAATTACAGTTTAGTATATACTGCAATTTAATTTGTAGCCTAATGCTGTTTGTATTAGGT 810
QY 186 -----GlyProTyrGlyGlyGlnGlnGln 194
DB 811 TTCAGAACGATCATAGTCTAATGTTTCTGTAGCTCCATATGGCAAAAGATTAGATT- 869
QY 195 TrpArgGlnGlnAsnAlaThr-----ProPro 203
DB 870 TATAAGCTAAAGGAGATGTTGCATAGTGTAGTAATGCTGTTTATACAGGAACACCT 929
QY 204 SerAspHisValValThrSerLeuProProProProLysAlaProSerProProArgGln 223
DB 930 CAAGATCATGTGTG---AATATGGTGGTCAAGAGGTGGGAATTTGGGGTCCACAGCAA 986
QY 224 ProProProProProProPheMetSerSerGlyGlySerAspTyrSerAsp 243
DB 987 CTTGTGCTGTGCTCT-----CACAGTGTGTTTCCACTTACCCGGT 1028
QY 244 ArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPhe 263
DB 1029 CGAACTGCTATACCTCACCTCAA---GCTCAACTCTTGTGTCAACCAAGCACTTTC 1085
QY 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGln 283
DB 1086 ACATACGATGAACCTGCTCCATTCGCAACAGAGGTTTCGCTCAGTCAAATTTCTAGGACAA 1145
QY 284 GlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLys 303
DB 1146 GGAGATTGGGTATGTTTATAAAGGAGTCTGCTAGTGCAAGAAAGTTGCGATGAAG 1205
QY 304 GlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIle 323
DB 1206 AGTCTTAACTTGGAAAGTGGACAGGGGAACGCGAGTTTCAAGCAGAGGTTGATATCAT 1265
QY 324 SerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343
DB 1266 AGCGGTGTCATCATCGTCTCTCTCTGTTGGATATTCATCTCTCGTGGTCAA 1325
QY 344 ArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisHis 361
DB 1326 AGACTTTGGTTATGAGTTTATACCTTAACAACACTTTTGAATTTTCATCTTCATGGTACA 1385
QY 361 ----- 361
DB 1386 TTCATCTAACAGATGTTTCTTGTATTAAACAAACCTTTTAAGTATGTTTCTCTTTAAT 1445
QY 362 -----GlyGluGlyArgProThrMetGluTrpSerThrArgLe 374
DB 1446 CAGGAACATGATTGAAATTTTCAGGAAGGGTCTGCGGTTTGGAGTTCGCTTACAAGAGT 1505
QY 374 uLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys----- 391
DB 1506 GAAGATTGCATTGGGATCAGCTAGAGGCTTGCAATTTTGCATGAGACTGTGAAGAAAT 1565

QY 391 ----- 391
Db 1566 CTTTATCTCACATATTTGCATCAGTTCTTCTATCTCGCTCTCTACAAATATTGAAAGATTGT 1625
QY 392 -----AsnProLysAlaIleHisArgAspIleLysAlaSerAs 404
Db 1626 ATATGTTACATCAATATTAGTGTACCTCTCGCATTTATCCACAGATATCAAGCTGCATA 1685
QY 404 nileLeuIleAspPheLysPheGluAla----- 413
Db 1686 CATTCTCTTGAATTCAGTTTGTAGAC-CAAGGTATGTGTATATATATCGACTCTTGTAC 1744
QY 413 ----- 413
Db 1745 TACTTTTACTTCTATCTCTCAATTTTGTTCCTCAATCTGTGTCGATGTGTATCAGT 1804
QY 414 -----LysValAlaAspPheGlyLeuAlaLysIleAlaSerAspTh 427
Db 1805 CTTATTGTGTAAATATATGCGGTGGCAGATTTTGGATTGGCTAAGCTATCTCAAGCAA 1864
QY 427 rAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
Db 1865 CTATACTCATGCTCTCCACTCGCGTCATCGGAACCTTTTGGGTAAAGCAGCTTGTGTAATGT 1924
QY 440 ----- 440
Db 1925 CTCACACTCACCACACTTATTAGTTCTTTTCACTTGTGTTTAACTATTTCTTGGATTCA 1984
QY 441 --TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheS 460
Db 1985 GATACTTAGCTCCAGATGATGATCATCAGCGGAAGTTATCCGACAAATCTGATGTTTCT 2044
QY 460 erPheGlyValValLeuLeuGluLeuIleThrGlyArgProValAspAlaAsnAsnV 480
Db 2045 CATTGAGTAATGCTCTTGGAGCTATAACCGAAGACCTCTCTCGATCTAACTGGA- 2103
QY 480 alTyrValAspAspSerLeuValAspTrp----- 489
Db 2104 --GAAATGGAAGATAGCTTGTGATGTTGGGTAAAGTCGGTCCCCGCCCTCTTCGGTTTACTT 2161
QY 489 ----- 489
Db 2162 GTTTAATCCCAAAACACTTTTCCAAGCAAAACAGAAACAATTTACTATTGTTGTGC 2221
QY 490 --AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspA 509
Db 2222 AGGCAAGGCCTTGTGTTGAAGCAGCTCAAGATGGAGATTACACCAATTTGCTCATC 2281
QY 509 lalysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaA 529
Db 2282 CAGCTCTAGAGCTTAAACTACAGTATCATCAAGAGATGGTTCAATGGCTTCTTGTGCAGTG 2341
QY 529 laCysValArgHisSerAlaArgArgProArgMetSerGlnIle----- 544
Db 2342 CAGCAATCAGACATTCAGCAAGAAGACGGCTTAGATGAGCCAGGT-TCAAAATACTCATA 2400
QY 544 ----- 544
Db 2401 CCACCTGTGTTCTATTGTTATATTTTACTCACAATTAATCTTGATGATAAATGTGAC 2460
QY 545 -----ValArgAlaLeuGluG 550
Db 2461 ATACTAATGAATCTTGAACATGTGTATGGTAAATGAAAGATTGTGACGACACTAGAAG 2520
QY 550 lyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrS 570
Db 2521 GAGATATGTCATGGATGATCTAGTGAGGGAACAGACCCAGGACAAAGCAGCTACTTGA 2580
QY 570 erSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysP 590
Db 2581 GCCCGGGAGGCGTGAGTCTCAGATGATGACGCAAGCTCGTACACGGCAGACATGAAAAAGT 2640
QY 590 heArgLysMetAlaLeuGlyThrGlnGlyTyrAsnAlaThrGlyGluTyrSerAsnProt 610

Db 2641 TCAGGAACCTGGCGGTAGAGAAATAAGAAATATCAACGACGT---GAATATGGTGA---A 2694
QY 610 hrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlyGlnThrThrArgGluM 630
Db 2695 CAAGTGAGTATGGCTTAAACCTTCTCTTCAAGTAGTGAA-----GAA 2739
QY 630 etGluMetGlyLysIleLysArgThrGlyGln 640
Db 2740 TGAATAGAGCTCAATGAAACGCAATCTCTCAG 2771
RESULT 13
AAF77095
ID AAF77095 standard; DNA; 2880 BP.
XX
AC AAF77095;
XX
DT 17-MAY-2001 (first entry)
XX
DE Arabidopsis gene #1.
XX
KW Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200114563-A1.
XX
PD 01-MAR-2001.
XX
PF 18-AUG-2000; 2000WO-CA000966.
XX
PR 19-AUG-1999; 99US-0149466P.
PR 13-OCT-1999; 99US-0159122P.
XX
PA (GORI/) GORING D.
PA (SILV/) SILVA N.
XX
PI Goring D, Silva N;
XX
DR WPI; 2001-244305/25.
XX
PT New proline-rich, extensin-like receptor kinase nucleic acids and
PT polypeptides useful for increasing plant wounding or pathogen resistance,
PT or for producing transgenic plants with increased wounding or pathogen
PT resistance.
PS Example; Fig 11; 91pp; English.
XX
CC The present invention relates to proline-rich extensin-like receptor
CC kinase (PERK). The PERK nucleic acids and polypeptides are useful for
CC increasing the resistance of plants to wounding and pathogens. These are
CC also useful for producing transgenic plants with increased wounding and
CC pathogen resistance compared with a wild type plant, as well as in assays
CC for identifying and developing compounds to inhibit and/or enhance
CC polypeptide function directly
XX
SQ Sequence 2880 BP; 814 A; 693 C; 569 G; 804 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.5e-27 Length: 2880
Score: 1209.50 Matches: 325
Percent Similarity: 43.85% Conservative: 92
Best Local Similarity: 34.17% Mismatches: 160
Query Match: 35.03% Indels: 376
DB: 4 Gaps: 26
US-10-086-464-2 (1-647) x AAF77095 (1-2880)
QY 2 SerSerAlaProSerProGlyThrGlySerProProSerProSerAsnSerThrThr 21
Db 98 AACTCAGCTCTCCACCTGTTGATTCCTCTCTCTCTAGTCACCCAGCTGATTCATCA 157
QY 22 Thr-----ThrProProAlaSer----- 28

Db	158		ACACCGCGCTGT	CAGAACCATCCAC	TCCCTCCAGATT	TCACAGCTTCCTCTTACCT	217
Qy	29	-----	-----	AlaProPro	ProThrThrProSer	ProProPro	40
Db	218	TCGATCTTCTCGG	TAAAGAGATCT	TCACACCTCCACCT	TTCGGATCTCTTCACCCGTT	277	
Qy	41	ProSerThr	ProThrSerProProSer	ProSerArgSerThr	ProSerAlaPro	---	59
Db	278	GATTCAC	---CCTTCTCGCGCC	CACCGCAGCT	CAACGAATCTCTCTCTCTCCAGAA	334	
Qy	60	-----	ProProSerProProThr	ProSerThrProSerThr	ProProProPro	---	74
Db	335	GATTCGAAAC	ACACCATCTCTCCACCAAT	TGAATCCAATGACAA	CAACACCTCTCCGCTCT	394	
Qy	75	-----	LeuProGlnProSerProProAla	-----	ProThrThrPro	86	
Db	395	CAGATCTTCAAT	CGCTCTCTCCATCGCTCGCGG	CCGGAATGTAGGACCC	CAACACCCG	454	
Qy	87	GlySerPro	-----AlaProValThr	ProThrArgAsnPro	-----	Pro	101
Db	455	GAATCAC	CCGTTACAAATCTCTCCAGCT	CCACGATCAGATCCT	CAAAATTCACCG	514	
Qy	102	ProSerVal	ProGlyProProSerAsnPro	-----	SerArgGluGlyGlySerPro	118	
Db	515	CCAGCTTCACCA	TTAGACCTTACCAATCTCTCCAGCT	CCACCAATACACCAT	CAGCAGCACCACT	574	
Qy	119	ArgProProSer	Pro-----SerProSerPro	-----	-----	129	
Db	575	TCTCTCCGCT	TAATCCCAACGCTCGCGG	CGCCATTCGCCACAGTAC	CAACCCAAACT	634	
Qy	129	-----	-----	-----	-----	129	
Db	635	CCTTCTAGT	GGACCTGTGGTCTCCATCTCTCACAT	CCCTCACATCCCTAGTAAAGAACTCCTACT	694		
Qy	130	-----	SerSerAspGlyLeuSerThrGly	-----	ValVal	139	
Db	695	CCAAACCAAG	CAATGAGATGGCGGTGGCGGT	GATCAAGGGAAGACTATG	754		
Qy	140	ValGlyLeu	AlaThrGlyValAlaLeuLeuVal	LeuValThrLeuLeuLeuCysLeuLeu	159		
Db	755	GTGGTATGCT	GTAGCGGTTTCGCAATGATGGCGCT	TATAGCGGTGTGTCTTAGTG	814		
Qy	160	CysLysLys	ValArgAlaGlyAspGluGluAspAlaTyr	-----	TyrValPro	175	
Db	815	AGAAGAAG	AAAGAGA-----AACAT	TGATAGCTATATCACTCACAGTACTTGCCA	868		
Qy	176	ProPro	-----	-----	177		
Db	869	CATCCCAAT	TCTCTGTTAAATCAGGTTTAA	AAATCTCACCTTATCTCTCTCTGATCAT	928		
Qy	177	-----	-----	-----	177		
Db	929	CTTCTATGT	GCTGAATCATCTCTGACTATCTTTGCT	TTTGTATGATGAGTACTTGA	988		
Qy	178	-----	ProProGlyProProLysAlaGlyGlyPro	TyrGlyGlyGlnGlnGln	194		
Db	989	TACGGTCAAG	ATCCAGTAAGNATCTCTCTGCT	CTAATGTTCAATGTATACAAAT	1048		
Qy	195	TrpArgGln	GlnAsnAlaThrProProSerAspHis	ValValThrSerLeuProProPro	214		
Db	1049	TCACAGCA	CAACAATCTCTATGGGAAACAGT	TTATGTCACAGCTGT	1096		
Qy	215	ProLysAla	ProSerProProArgGlnProProPro	ProProProProPheMetSer	234		
Db	1097	-----	GGTGGTTATCTCTCATCA	CA-----	ATGCAA	1123	
Qy	235	SerSerGly	GlySerAspTyrSerAspArgPro	ValLeuProProProSerProGlyLeu	254		
Db	1124	TCAAGTGG	CACCTGAC-----	-----	TCGTCT	1147	
Qy	255	ValLeuGly	PheSerLysSerThrPheThrTyrGlu	GluLeuAlaArgAlaThrAsnGly	274		


```
QY 130 -----SerSerAspGlyLeuSerThrGly-----ValVal 139
D 695 CCAACCAAGCAATGGAGATGGCGGTGGCGGTATCAAGGGAAGACTATG 754
QY 140 ValGlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeu 159
D 755 GTTGTATGGCTGTAGCGGTTTCGAATCTGCGGCTTATAGGCGTTGTGTTCTAGT 814
QY 160 CysLysLysLysArgArgAspGluGluAspAlaTyr-----TyrValPro 175
D 815 AGAAGAAAGAAAGAGA-----AACATTGATAGCTATAATCACTCAGTACTTGCCA 868
QY 176 ProPro----- 177
D 869 CATCCCAATTTCTCTGTTAAATCAGGTTTAAAAATCTCACCTTTATCTCTCTCATCAT 928
QY 177 ----- 177
D 929 CTTCATGTGCTTGAATCATCTCTCTGACTATCTTTGCTTTTGTATGATAGATGATCTTA 988
QY 178 -----ProProGlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGln 194
D 989 TACGCTCAAGATCCAGGTAAGGATACTCTCTGCTGCTCTTAATGTTCAATGTATTAACAAT 1048
QY 195 TrpArgGlnGlnAsnAlaThrProSerAspHisValValThrSerLeuProPro 214
D 1049 TCAGAGCAACAACATCTCTATGGGAACAGATTATGTCAGCTGGT----- 1096
QY 215 ProLysAlaProSerProProArgGlnProProProProProProPheMetSer 234
D 1097 -----GGTGGTTATCTCATCATCAA-----ATGCAA 1123
QY 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProSerProGlyLeu 254
D 1124 TCAAGTGGCACACCTGAC-----TCTGCT 1147
QY 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274
D 1148 ATACTCGGAAGTGGCCAGACTCATTTAGTTACGAAGAGCTTGCTGAGATAACACAGGC 1207
QY 275 PheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeu 294
D 1208 TTTGTCTCGCAAAACATTTCTTGAGAGGCGGATTTGGATGTGTCTATAAAGGTACATTTG 1267
QY 295 ProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArg 314
D 1268 CAGGATGGTAAAGTTGTTCGGTTAAGCAGCTTAAAGCTGGAAGTGGACAGGTGACCGT 1327
QY 315 GluPheGlnAlaGluValGluIleIleSerArgValHisArgHisLeuValSerLeu 334
D 1328 GAATTCAAAGCAGAGGTTGAGATCATCAGCCGGTTTCATCATCGCCATTTGTTCTCTCTG 1387
QY 335 ValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsn 354
D 1388 GTTGTACTGCATTTTCAGACCAGCATAGATTGCTTATCTATGATATGTTTCTAATCAA 1447
QY 355 AsnLeuGluLeuHisLeuHisGlyGlu----- 363
D 1448 ACCTTGGAGCATCATTTGCATGGTGAGTGACTTTGTACCATTTCGTTATAGATTAAGACT 1507
QY 364 -----GlyArgProThr 367
D 1508 TTTTITTTAGCTTACGTGTAGACTGACTCGCTTTACGCTTTTAGGAAAGGTTTGCCAGT 1567
QY 367 rMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrIle 387
D 1568 TTTTAGAGTGGTCTAAGAGAGTCCGGATCGCTATAGGATCAGGCAAAAGGTTTGCCATATCT 1627
QY 387 uHisGluAspCysAsn----- 392
D 1628 TCACGAAGACTGTAA- GTAATGCTTCACATTTTCTTAGTTGTGTGCTTTGTTATGCAAC 1686
```

```
QY 393 -----Proly 394
D 1687 TTCATAGTTTAAACAGAAAGCAAAAATCATATCTCTGTTGTTTATTATTACAGGTCAATCCGAA 1746
QY 394 sIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLys 414
D 1747 AATCATTCACAGAGATATAAAGTCAGCAAAATATTCTTCTAGATGATGAATGAAGCTCA 1806
QY 414 8----- 414
D 1807 -GGCAATTAATGAATCCCTCTTTTCGTTAAATCTATCTTATGACTGCTAAAGTTTAGTTA 1865
QY 415 -----ValAlaAspPheGlyLeuAlaLysIleAl 424
D 1866 ATGAGACTTGTCTGTTTTTTTGGGATGTTTAGTGTGCTGATTTTGGACTTGTCTAGACTCAA 1925
QY 424 aSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
D 1926 TGATACAACACAAAACCTCATGTTTCAACTCGGTTATGGAAACCTTCGGGTAAAGCAACAT 1985
QY 440 ----- 440
D 1986 TCATCACAAACTCTACTCCAAAACCTGGACCTTATTGTATCCATGCCTGATGAAAGTTTG 2045
QY 440 ----- 440
D 2046 TTATATATGGCTTGAGGCAACAAATTTGGATCAAACTGAAATCTTTTATTGATCGTATGGCT 2105
QY 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 453
D 2106 GCATGACATGTTTTGTGTTAAGGTACCTAGGCGCGGAATATGCAATCAAGTGAAATATGA 2165
QY 453 hrGluLysSerAspValPheSerPheGlyValValLeuLeuLeuLeuIleThrGlyArg 473
D 2166 CTGATAGATCCGATGATATCTCATTCGSGGTGTTCTCTTAGAGCTTGTAACTGACGGA 2225
QY 473 rgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrp----- 489
D 2226 AACAGTTTGACACAGACTCAGCCTCTAGGAGAAGAGAGTTTGGTTGAATCGGTAAGAATCC 2285
QY 489 ----- 489
D 2286 AACTTTCAACAATCTTCAATAATAGTAGAATGGCCCTAGTATATCTATATAGTACTTA 2345
QY 490 -----AlaArgProLeuLeuAsnArgAlaSerGluGlnGlnGlyAspPheGlu 504
D 2346 TAAATGAATCAACAGCGCGCGCTGCTTCTCAAAGCCATTGAGACCGGAGATTTAAGC 2405
QY 505 GlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetVal 524
D 2406 GAACTGATTGATACACGGCTTGAAAAAGCGTTATGTGGAGCATGAAAGTCTTCAGAATGATC 2465
QY 525 AlaCysAlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGln--- 543
D 2466 GAGACAGCGCTGTCATGTTGTAGACATTTCTGTGTCCAAAACGCCACCGCATGTTTCAAGTA 2525
QY 543 ----- 543
D 2526 ATTCTGACTAACCAAAAGTCCAAAGCTCCCATATATAGTAAACAGTATTTCTCATCT 2585
QY 544 -----IleValArgAlaLeuGluGlyAsnValSerLeu 554
D 2586 GAAAACTTATCTACTCTTCGAAATAAGGTTGTGAGAGCATTTGGACTGCGAGAGACTCG 2645
QY 555 SerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGly 574
D 2646 GGAGATATTAGCAACGGAATCAAAATTGGGCAA----- 2678
QY 575 SerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAla 594
D 2679 AGCACAACTTATGACTCAGGCAATACAATGAAGACATTTATGAATTCAGGAAAAATGGCG 2738
QY 595 LeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGly 614
```

[illegible]

1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450</
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

Db	1102	GGCGGTAATGCTGGTACGAACAGGCACATGTTATCACA---ATGCCACCAACAATCCA	1158
Qy	216	sAlaProSerProProArgGlnProProProProProProPheMetSerSerSe	236
Db	1159	TGCTAAA-----TATATATCTAG	1176
Qy	236	rGlyGlySerAspTyrSerAspArgProValLeuProProProSerProGlyLeuValle	256
Db	1177	TGGAGGTGTGTACGAAGGAGAAC-----AATTCGTTCGCAAAAACAT	1221
Qy	256	uGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSe	276
Db	1222	TTCAATGCCATCTCGAATGTTCTCCTACGAAGAACTTTCAAAAGCAACTCGTGGATTTTC	1281
Qy	276	rGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProse	296
Db	1282	AGAGGAAACCTTTTGGAGAGGCGGTTTCGGATATGTTCCAAAGGAGTGTGTAAAA	1341
Qy	296	rGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPh	316
Db	1342	CGGCACAGAAGTTGCGGTGAAGCAGCTGGAAGATTGGGAGCTATCAAGGGGAAAGAGAA	1401
Qy	316	eGlnAlaGluValGluIleIleSerArgValHisHisArgHisLeuValSerLeuValGl	336
Db	1402	CCAAGCTGAGGTGCACAACTCAGTAGGGTTCATCAAGCAGCTCGTTTCATTTGGTTGG	1461
Qy	336	YTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLe	356
Db	1462	TTATTGGCTTAATGGAGATAAAGACCTCTGGTTTACGAGTTGTTCCTTAAGATACCTT	1521
Qy	356	uGluLeuHisLeuHisGly-----	362
Db	1522	GGAGTTCCACTTGCATGGTAAATAGATATATGATTTTCATCCTTTTTTGATTTTGTCTTTA	1581
Qy	363	-----GluGlyArg--ProThrMetGl	369
Db	1582	GTTCAATTATGTTGAGTATTGTGAGATATATGTGTTGAGAGAACAGAGGAGCGTGTGGA	1641
Qy	369	uTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGl	389
Db	1642	ATGGAAATCAGGCTCAGGATTGCTGTAGGAGCAGCAAAAGGATTAGCTTATCTTCATGA	1701
Qy	389	uAspCys-----	391
Db	1702	GGATTGTGAGTGTCTCTTCATAATGGGAATGACAATGCCCTTTTTTGGGTTTTTCGAC	1761
Qy	392	-----AsnProLysIleIleHisArgAspIl	400
Db	1762	ACTGATATTGATTTCTGGTCTGTAATGAGGAGCTCCAACATAATTCACCGTGATAT	1821
Qy	400	eLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLys-----	414
Db	1822	CAAAAGCAGCTAATATCTCTCTAGATTCCAAATTTGAGGCCAAAGGTGATCTGCTCTTAAT	1881
Qy	414	-----	414
Db	1882	CTTATCAAAAGTTGGTTTTTGTAGAACAGAGTTTGCCCATTTTCGGTTTATAGTCACACCAT	1941
Qy	415	-----ValAlaAs	417
Db	1942	TTTGTCTTATTAAACAGATTACATGTAGCTTAGAGTTCTTGTTCTCTTCAGGTCTCTGA	2001
Qy	417	pPheGlyLeuAlaLysIleAlaSerAspThrAsn-----ThrHisValSerThrAr	434
Db	2002	CTTTGGACTAGCCAAAGTTTTTCTCAGACACCAATTCATCATTCATCATATCTACTCTCT	2061
Qy	434	gValMetGlyThrPheGly-----	440
Db	2062	AGTGGTAGGAACCTTCGGGTAAACACCATCCATGCCTTTATATGTTGTGTGTCATTGT	2121
Qy	440	-----	440

2122	Db	GTTTAAAAATATTAATTTACGGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTTCTTCCTT	2181
440	Qy	-----	440
2182	Db	GTTGTATCGTAAATCCAGATCAAAAGATTATTGATTACTAAATGCCTTTGTGCACCGTCT	2241
441	Qy	-----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuTh	453
2242	Db	ATTTGGTATGCATTAAACAGATACATGGCTCCAGAAATACGGCTCCAGTGGTAAAGTAAC	2301
453	Qy	rGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgAr	473
2302	Db	TGATAAATCAGATGATATTCCTTTGGGGTGGCTTCTAGAACTCATCACTGGACGCTCC	2361
473	Qy	gProValaspAlaAsnAsnValTyrValaspSerSerLeuValaspTrp	489
2362	Db	ATCAATTTTCGCCCAAGATTCTTCCACAAACCAGAGTTTAGTAGACTGGGTAAAGTCAAAG	2421
489	Qy	-----	489
2422	Db	TACATGATGATGATGATACCATTAGTGTTCCTTTTCCCCTAGTTATTATAGATGAAA	2481
489	Qy	-----	489
2482	Db	ATGAAACTCCAGTATAACTCGAAGCTTGAACCTTATTAGACTGTAAATTTTGATTATTC	2541
490	Qy	-----AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaA	508
2542	Db	CTCAGGCGAGGCCATTGCTTACGAAAGCAATCTCTGAAGAAAGTTTGTGACTTCTTCTGTAG	2601
508	Qy	spAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaA	528
2602	Db	ACTCAAGGTTGGAGAGAAATTTACGATACAACTCAGATGGCAACAATCGCTGTGTGCTG	2661
528	Qy	laAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIle	544
2662	Db	CTGCTTGATACGCCAATCAGCTGGCTTCGGCTTAGAATGAGCCAGGT-CTGAGATTAA	2720
544	Qy	-----	544
2721	Db	GTTTAAATACATGTATTTCGTCCATATATCGAAGAGGAAACGCTAATCCATCAATACAT	2780
545	Qy	-----ValArgAlaLeuGluGlyAsnValSerLeuS	555
2781	Db	TTATATCTTGAAAAAACTTGAATAGGTAGTACGTCTCTTGAAGGCGAGGTGGCCCTGA	2840
555	Qy	erAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlyS	575
2841	Db	GAAAGGTCGAAG------ACTGGGAATAGCGTACCTATAGCTCTTCTGAAACCC	2891
575	Qy	erThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaL	595
2892	Db	CGAATGCATC-ACACCCAGGTATGGAACAATAAGAGGAGATTCTC-----	2935
595	Qy	euGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyL	615
2936	Db	-----GACACAGGTTCAACCGATGGTTACACTTCAAGATATCGAG	2975
615	Qy	euTyrProSerGlySerSerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysI	635
2976	Db	TTAACCTTCTCAGTCGAGCAGTGAACATCAACACAGGTGAATACTAGTTCCACAGGTTCAA	3035
635	Qy	leLysArgThrGlyGlnGlyTyrSerGlyProSerLeu	647
3036	Db	TA-----GGGCAAGTTTCCACCAATTA	3058

Search completed: April 24, 2004, 22:22:00
Job time : 666 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 24, 2004, 20:30:41 ; Search time 133 Seconds
(without alignments)
2699.648 Million cell updates/sec

Title: US-10-086-464-2
Perfect score: 3453
Sequence: 1 MSAAPSGTGPSPSPNST.....REMEMGKIKTGQYSGPSL 647

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10086464/runat_23042004.083016.5642/app.query.fasta_1.839
-DB=Issued_Patents_NA -QPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsm2 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10086464 @CGN 1 1.85 @runat_23042004.083016.5642 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	756	21.9	3239	4	US-09-228-986-9
2	744.5	21.6	1593	4	US-09-602-472A-1
3	740	21.4	2868	4	US-09-228-986-4
4	712	20.6	1488	4	US-09-579-182-4
5	703.5	20.4	2513	4	US-09-228-986-13
6	678	19.6	2432	4	US-09-228-986-7
7	667.5	19.3	1559	4	US-09-602-472A-3
8	604.5	17.5	4104	3	US-08-881-706-1
9	602	17.4	1224	4	US-09-579-182-3
10	600.5	17.4	3097	4	US-09-228-986-11
11	581.5	16.8	3590	1	US-08-587-889-1
12	581.5	16.8	3590	4	US-09-016-434-1093

13	581.5	16.8	3590	5	PCT-US96-09193-1	Sequence 1, Appli
14	578	16.7	2114	4	US-09-602-472A-5	Sequence 5, Appli
15	575	16.7	2749	1	US-08-265-628-1	Sequence 1, Appli
16	572	16.6	2571	1	US-07-717-331F-9	Sequence 9, Appli
17	572	16.6	2833	1	US-07-717-331F-1	Sequence 1, Appli
18	567	16.4	1554	2	US-08-587-680A-24	Sequence 24, Appli
19	566.5	16.4	2389	4	US-09-228-986-1	Sequence 1, Appli
20	562	16.3	2336	4	US-09-228-986-10	Sequence 10, Appli
21	559.5	16.2	2638	4	US-09-228-986-8	Sequence 8, Appli
22	546	15.8	2749	1	US-07-717-331F-4	Sequence 4, Appli
23	545	15.8	966	1	US-08-447-185-2	Sequence 2, Appli
24	545	15.8	2443	1	US-08-447-185-3	Sequence 3, Appli
25	543	15.7	2686	4	US-09-228-986-3	Sequence 3, Appli
26	531.5	15.4	5733	2	US-08-473-553A-1	Sequence 1, Appli
27	486	14.1	2943	4	US-09-503-922-2	Sequence 2, Appli
28	460.5	13.3	2649	4	US-09-228-986-12	Sequence 12, Appli
29	430	12.5	2568	4	US-09-228-986-2	Sequence 2, Appli
30	401	11.6	1926	4	US-09-249-585A-2	Sequence 2, Appli
31	401	11.6	1926	4	US-09-410-399-3	Sequence 3, Appli
32	401	11.6	2580	3	US-09-050-863-2	Sequence 2, Appli
33	401	11.6	2580	4	US-09-359-081-2	Sequence 1, Appli
34	401	11.6	5452	2	US-09-130-114-1	Sequence 1, Appli
35	401	11.6	8705	4	US-09-647-344A-14	Sequence 14, Appli
36	401	11.6	9600	3	US-08-910-647-1	Sequence 1, Appli
37	401	11.6	9600	4	US-09-620-925-1	Sequence 1, Appli
38	401	11.6	10596	1	US-07-884-811-15	Sequence 15, Appli
39	401	11.6	10596	1	US-07-885-971-15	Sequence 15, Appli
40	401	11.6	10596	1	US-08-087-783A-15	Sequence 15, Appli
41	401	11.6	10596	1	US-08-194-088B-15	Sequence 15, Appli
42	401	11.6	10596	2	US-08-194-087-15	Sequence 15, Appli
43	401	11.6	10596	5	PCT-US93-04648-15	Sequence 15, Appli
44	401	11.6	16080	4	US-09-724-566A-48	Sequence 48, Appli
45	395.5	11.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-228-986-9

; Sequence 9, Application US/09228986

; Patent No. 6359198

; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Niels

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling

; FILE REFERENCE: 11000/1020

; CURRENT APPLICATION NUMBER: US/09/228,986

; CURRENT FILING DATE: 1999-01-12

; NUMBER OF SEQ ID NOS: 130

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 9

; LENGTH: 3239

; TYPE: DNA

; ORGANISM: Eucalyptus grandis

US-09-228-986-9

Alignment Scores:

Pred. No.:	1.88e-23	Length:	3239
Score:	756.00	Matches:	190
Percent Similarity:	51.93%	Conservative:	92
Best Local Similarity:	34.99%	Mismatches:	132
Query Match:	21.89%	Indels:	130
DB:	4	Gaps:	17

US-10-086-464-2 (1-647) x US-09-228-986-9 (1-3239)

Qy 89 ProProAlaProVal-ThrProThrArgAsnProProSerValProGlyProPr 108

Db 1759 CTTCCCAACCAATTGCGCCCTATTACTTCATCGCTCCCTATGCTTTT-----CC 1811

Qy 108 oSerAsnProSerArgGluGlySerProArgProProSerSerProProSe 128

```
Db 1812 AGACAAT-----GGAGGAACC----- 1827
Qy 128 rProSerSerAspGlyLeuSerThrGlyValValGlyIleAlaIleGlyValAl 148
Db 1828 -----GCCATAAGCAAAAGGTGTATAGTGTGGATCGCTATTGGCGGCACCGT 1874
Qy 148 aleuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysArgArgAspG1 168
Db 1875 TCTGGTCTTGCCCTTGTTGATTAGGTATTATCTATTCCACAAAGAAACGGCGGA 1934
Qy 168 uGluAspAlaTyrTyrValProProPro-----ProProGlyProLysAl 184
Db 1935 GAAA---GCTCTCGAGTTCAGCAGACCCCTTCGCATCTCGGCACCCAGTGGGAAAGATAG 1991
Qy 184 eGlyGlyProTyrGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSe 204
Db 1992 CGGAGAGCGCCACAACTGAAGAGGACGATGG----- 2025
Qy 204 rAspHisValValThrSerLeuProProProLysAlaProSerProProArgGlnPr 224
Db 2025 ----- 2025
Qy 224 oProProProProProPheMetSerSerSerGlyGlySerAspTyrSerAspAr 244
Db 2025 ----- 2025
Qy 244 gProValLeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheTh 264
Db 2026 -----TTCTC 2030
Qy 264 rTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnG1 284
Db 2031 CTATGATGAACCTTAAGAGGTGCACCAATAATTCTCCGATACCAATGAATTAGGCTTCGG 2090
Qy 284 yGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysG1 304
Db 2091 AGGATACGGAAAGGTGTACAGGGAGTCTTCTCGATGCTCATATATAGCAATCAAAAG 2150
Qy 304 nLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSe 324
Db 2151 ACCTCAGCAGGGGTGATGCGAGGTGCACCCAGGTTCAGACAGAAATCGAGCTGCTTTC 2210
Qy 324 rArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysAr 344
Db 2211 GCGGGTTCATCAACAAGATCTTGTTGGCTCATAGGATTCGTTTCGAGCAAGGAGACA 2270
Qy 344 gLeuLeuValTyrGluPheValProAsnAsnLeuLeuGluLeuHisLeuHisGlyGluG1 364
Db 2271 GATGTGGTCTATGAATATATGCTTAACGGGACGCTCAGGGATAGCTTTGACAGGAAATC 2330
Qy 364 yArgProThrMetGluTTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLe 384
Db 2331 AGGCATTTATCTTGATTTGAAGAGAGGCTTCGTATAGCTCTAGGTTGCGCTAGAGACT 2390
Qy 384 uSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAs 404
Db 2391 AGCTTATCTGCAGCAACTCGCAATCTCCATTTATCCACAGAGATGTCAAGTCCACCAA 2450
Qy 404 nIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAl 424
Db 2451 TATCTTGTGGACGAACATCTGACGCCCAAAAGTCGGGATTTTCGGTTTGTCCAAACTGTG 2510
Qy 424 aSerAspThrAsnThr---HisValSerThrArgValMetGlyThrPheGlyTyrLeuAl 443
Db 2511 ATCGACACGGGAGGGGCAGCTTTCGACGCAAGTGAAAGCAGCTGGGCTATTGGA 2570
Qy 443 aProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVa 463
Db 2571 TCCCGAATACTACATGAGTCAACAGCTGACAGAAAGAGCGATGTGTACAGCTTCGGGT 2630
Qy 463 lValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnValTyrValAs 483
Db 2631 GGTCTATCTTGAGCTCATCATGCAAGCAACCGATTTGAGAGAGGGCAAGTATGTCTGTCG 2690
```

```
Qy 483 pAspSerLeuValAspTTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPh 503
Db 2691 CGAG-----ATTCGACCCCATGGACAGAACGACAG-----GACTA 2729
Qy 503 eGluGly-----LeuAlaAspAlaLysMetAsnAsn---GlyTyrAspArgGluG1 519
Db 2730 CTACGGCGTGAAGGAAATGATGGACCCGCTCCATGAGGAGCATGGCTAC---CTCGTCGG 2786
Qy 519 uMetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgArgPr 539
Db 2787 GTTCAGCAGGTTCCTGGATTTCGGCATGCGATGTGTGAGGAGTCGGCTCGGACCCGCC 2846
Qy 539 oArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnG1 559
Db 2847 CACATGAGCGAGGTGTGAAGCGATCGACCATGTTGCAG-----AACGA 2894
Qy 559 uGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAs 579
Db 2895 TGGGATACACCAACTCGAGTCGGCATCGCTCGTCG-----GCGACGGACTTTGG 2945
Qy 579 pSerSerGln-----TyrAsnGluAspMetLysLysPheArgLysMe 593
Db 2946 GTCGACGAAGGCGCTCCTCGCATCCGTACACGATGCCCTTACCAG----- 2994
Qy 593 tAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTy 613
Db 2995 -----AAGGAAGTTAGCTATAGGATTCCTTTGATTATAGT-----GGTGGATA 3038

RESULT 2
US-09-602-472A-1
; Sequence 1, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
US-09-602-472A-1

Alignment Scores:
Pred. No.: 2,9e-23 Length: 1593
Score: 744.50 Matches: 172
Percent Similarity: 57.07% Conservative: 66
Best Local Similarity: 41.25% Mismatches: 116
Query Match: 21.56% Indels: 63
DB: 4 Gaps: 9

US-10-086-464-2 (1-647) x US-09-602-472A-1 (1-1593)
Qy 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProSerProGlyLeu 254
Db 181 TCGTCGGGTTCGAAAGACGGGTTCGAGGTCAGATT----- 216
Qy 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274
Db 217 -----GCTGCTCATACATTCTTTCGGGAGCTTGACGTGCAACAAACAAT 264
```

Qy 588 LysLysPheArgLysMetalLeuGlyThrClnGlutTyraenAlaThrcGly 604
Db 1333 GAAAAAGACGACAAGAATGGCAAGGCAGCTTTGTATGGGAATGCTAACCGT 1383

RESULT 3
US-09-228-986-4.
; Sequence 4, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 2868
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-4

Alignment Scores:
Pred. No.: 7,71e-23 Length: 2868
Score: 740.00 Matches: 182
Percent Similarity: 50.91% Conservative: 69
Best Local Similarity: 36.92% Mismatches: 133
Query Match: 21.43% Indels: 109
DB: 4 Gaps: 13

US-10-086-464-2 (1-647) x US-09-228-986-4 (1-2868)

Qy 90 ProLaProValThrProProThrArgAsnProProSerValProGlyProPro--- 108
Db 858 CCATGTCGGGATCTCTCCCTTTTCTCCTCCTCCTCGTTGTACCACCCGCCAGTA 917
Qy 109 -----SerAnProSerArgGluGlySerProArgProProSerSerProSerPro 126
Db 918 GCTGGAGCATGGAGCAGCGGTGCAGAGCTCTCCC----- 953
Qy 127 ProSerProSerSerAspGlyLeuSerThrGlyValValGlyVileAlalleGlyGly 146
Db 954 -----AGCACAGAGCCATTGCTGGGGAGTAGTCTGCAGGT 989
Qy 147 ValAlaLeuLeuValIleValThrLeulleCysLeuLeuCysLysLysLysArgargarg 166
Db 990 GCTGCCCTCTTATTT----- 1004
Qy 167 AspGluGluAspAlaTyTyrrValProProProProGlyProLysAlaGlyGly 186
Db 1005 -----GCAGTCCTGCAATTGGGTTC 1025
Qy 187 ProTyGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAsphis 206
Db 1026 GCCTGG-----TGCGTCGCAGAAAG-----CCACAGAGCAC 1058
Qy 207 ValValThrSerLeuProProProProLysAlaProSerProProArgGlnProPro 226
Db 1058 ----- 1058
Qy 227 ProProProProPhMetSerSerSerGlyGlySerAspTyrrSerAspArgProVal 246
Db 1059 -----TCTTTTGATGACTCTGCT 1076
Qy 247 LeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyGlu 266
Db 1077 -----GAGGAGGATCCAGAGTTCACTTAGCCCACTTAG--AGGTTCTCATTCAGG 1127
Qy 267 GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyclnGlyPhe 286
Db 1128 GAATTCAGGTTGCAACTGATGGTTTTAGCAATAGAAAACATTTCTGGCAGAGGTGGTTT 1187

```
QY 207 GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys 306
DB 1188 GGAAGGGTGAACAAGGCGCTTCACAGATGTTCTCGTGGCTGTAAACGTCGAG 1247
QY 307 ---ValGlySerGlyGlnGlyLysGluValGlnAlaGluValGluLeuLeuSerArg 325
DB 1248 GAAGAGCGTACACCGGGTGGAGAGTTCAGTTTCAACAGAGTGGAGATGATGAAGCATG 1307
QY 326 ValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeu 345
DB 1308 GCAGTACATAGAACCTCTTCGATCTAGTGTGATTTCTGCATGACACCCCACTGAACGGCTG 1367
QY 346 LeuValTyrGluPheValProAsnAsnLeuGlu-----LeuHisLeuHisGlyGlu 363
DB 1368 CTGTGTTATCCCTACATGCCCAATGGAAGTGTGCTTCATGCTACGAGAGGGGCCAA 1427
QY 364 GlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGly 383
DB 1428 AATGACCCACCTTAGATTGGCCAACTCGCAAGCGCATAGCATTTGGTTCGCAAGAGGT 1487
QY 384 LeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSer 403
DB 1488 CTCCTCTACTGATGATCATTTGTCATCTAGATTAATTCACCGGATGTCGAAGCTGCT 1547
QY 404 AsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIle 423
DB 1548 AACATCTTACTGGATGAAGAATATAGGCAGTGTGGGGGATTTTGGCTTGGCAAACTT 1607
QY 424 AlaSerAspThrAsnHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla 443
DB 1608 ATGGATTATAAGACACACATGTTACGACGGCTGTTCTGGTGAACCATTTGGCCACATAGCA 1667
QY 444 ProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVal 463
DB 1668 COTGAGTACCTTCTACTGGAAAGTCTTCGGAAGAGACAGACGATATTTGGATATGGAATC 1727
QY 464 ValLeuLeuGluLeuIleThrGlyArgProValAspAlaAsnValTyrValAsp 483
DB 1728 ATGTTGCTGGAACTATTATACGGACACACGGGCATTTGACCTTGACCGTTTACCAATGAT 1787
QY 484 AspSer-----LeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGly 501
DB 1788 GATGATGTCATGTTGCTTGAAGTAAAGGCTTACTA-----AAAGAGAGA 1835
QY 502 AspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAla 521
DB 1836 AGGCTTGATATGCTAGTTGATCTGATCTTAAGAACAAATTATGTTGAAGCAGAGGTGGA 1895
QY 522 ArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMet 541
DB 1896 CAACCTATTCAAGTTGCATTACTTTGTACAAAGGGTACCAATGGATAGACCAAGATG 1955
QY 542 SerGlnIleValArgAlaLeuGluGlyAsn-ValSerLeuSerAspLeuAsnGluGlyMet 561
DB 1956 TCTGAAGTGGTAAGATGTTGAAGGGATGCTTAGCTGAGAGATGGGAGGAATGGCAA 2015
QY 561 ArgProGlyGlnSerAsnValTyrSerSerTyrGly 573
DB 2016 AAGG-----TGAAGTCTGACGCA 2034

RESULT 4
US-09-579-182-4
; Sequence 4, Application US/09579182
; Patent No. 6500628
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; FILE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREOF
; FILE REFERENCE: MNI-161
; CURRENT APPLICATION NUMBER: US/09/579,182
; CURRENT FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 7
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-579-182-4

Alignment Scores:
Pred. No.: 6,04e-22 Length: 1488
Score: 712.00 Matches: 178
Percent Similarity: 53.07% Conservative: 73
Best Local Similarity: 37.63% Mismatches: 160
Query Match: 20.62% Indels: 62
DB: 4 Gaps: 14

US-10-086-464-2 (1-647) x US-09-579-182-4 (1-1488)

QY 205 AspHisValValThrSerLeuProPro-----ProProLysAlaProSer 219
DB 22 GAACACACGATAAATCCCTTCGCTCCGCTATATTTCTTAAATCCGATAAATCAAGTCT 81
QY 220 ProProArgGlnProProProProProProPheMetSerSerSerGlyGlySer 239
DB 82 -----CGACAGGTCCT----- 93
QY 240 AspTyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSer 259
DB 94 -----GNAACAGAGAGGAGCTAACTGCTCAAAAGAGGCGCTACTGCGCATATTGCT 147
QY 260 LysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
DB 148 GCACAAACCTTTACTTTCCGAGAGTGTAGTCCGCCCACTAAAAAATTTTCGACCGGAATGT 207
QY 280 LeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeu---ProSerGlyLys 298
DB 208 CTTCTTGAGAGAGGAGGTTTCGGACGCTTTTACAAAGGTCGTCTAGAGACCACAGAGCAG 267
QY 299 GluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAla 318
DB 268 ATAGTAGCTGTAAACAGCTTGATCGAAACGCTCTACAAGGAAACACAGAGAGTTTCTTGTA 327
QY 319 GluValGluIleLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCys 338
DB 328 GAGGTTCTTATGCTGAGCGCTTCTGCATCATCCCACTTTGTGAATTTGATGGTTATTGT 387
QY 339 IleAlaGlyAlaLysArgLeuValTyrGluPheValProAsnAsnLeuGluLeu 358
DB 388 GCTGATGGGACCCAGCGTCTTCTTGTGATGAGTATATGCCACTAGGATCATGGAGGAT 447
QY 359 HisLeuHisGly-----GluGlyArgProThrMetGluTrpSerThrArgLeuLysIle 376
DB 448 CATCTACACGATCTTCCACCTGATAAAGAGCGCTCTAGACTGGAGTACTAGATGACAATA 507
QY 377 AlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleLeu 396
DB 508 CGCGCAGGAGCAGCAAAAGGAGCTGGAGTATCTGCATGATAAAAGCGAATCCGCTGTGATC 567
QY 397 HisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLysValAla 416
DB 568 TACAGAGACCTGAAATCATCCAAACATTTCTTCGGTGATGGCTATCACCCAAAGTTATCT 627
QY 417 AspPheGlyLeuAlaLysIleAlaSer---AspThrAsnThrHisValSerThrArgVal 435
DB 628 GATTTGGTTAGCTAAGTTAGGTCCTCGGCGCATATAAACACATGTGTCAACTCGTGTG 687
QY 436 MetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLys 455
DB 688 ATGGCACATATGGTTATTGTGCACCGAATATGCCATGACAGGGCAACTCACTGATGAA 747
QY 456 SerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgProVal 475
DB 748 TCCGATGTTTATAGCTTTGGGGTGTGTCTTCGAGCTCATCCCGGTCGAAAGCTATT 807
```


444	Qy	ProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVal	463
1794	Db	CCTGAAATATGTCATGTCAGGAATCTATACAAATGAAGAGCCATGTTATAGTCTTTGGAGTC	1853
464	Qy	ValLeuLeuGluLeuIleThrGlyAArgProValAspAlaAlaAsnAsnValTyrValAsp	483
1854	Db	GTAATGCTTGAGCTTTTGACAGGCGGAGCCCTTGGATAGTTCAGAAACAAGGTCAGAA	1913
484	Qy	AspSerLeuValAspTrpAlaAArgProLeuLeuAsnAArgAlaSerGluGlnGlyAspPhe	503
1914	Db	CAGTCATTAGTAAGTGGGCCACCCCTCAGCTTCAT-----GATATT	1955
504	Qy	GluGlyLeuAla-----AspAlaLysMetAsnAsnGlyTyrAspAArgGluGluMet	520
1956	Db	GATGCATTGGCAAGATGGTTGATCCAGCTTTTGAAGGGAAGTATCTCGTCAAAATCTCTC	2015
521	Qy	AlaAArgMetValAlaCybAlaAlaAlaCysValAArgHisSerAlaAArgAArgProAArg	540
2016	Db	TCGGCTTTTGCATATATTATGGCCCTCTGCATTCAGCCCTCAACAGAAATTCCTCTCCA	2075
541	Qy	MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerSerAspLeuAsnGluGly	560
2076	Db	ATGCTCTGAAGTGTGCAAGCATTTGGTTGCTGATGTCACCGGGCTACGCTCAATAAGAGG	2135
561	Qy	MetAArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer	580
2136	Db	ATGACAGGAGATGNA-----ACTGCAGACCCAGCATCTCT	2168
581	Qy	SerGlnTyr	583
2169	Db	GCAGATTAT	2177

RESULT 6

```

US-09-228-986-7
; Sequence 7, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1993-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2432
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-7

```

Alignment Scores:

Pred. No.:	2.44e-20	Length:	2432
Score:	678.00	Matches:	147
Percent Similarity:	63.0%	Conservative:	68
Best Local Similarity:	43.1%	Mismatches:	104
Query Match:	19.6%	Indels:	22
DB:	4	Gaps:	8

US-10-086-464-2 (1-647) x US-09-228-986-7 (1-2432)

Qy	254	LeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsn	273
Db	1046	ATTTGCTCTGGGCAATTTGAAG--AGATCTCATTTCAAAGAGTTACGTGTTTCCACTAAT	1102
Qy	274	GlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyVal	293
Db	1103	AAATTTTAGTAGCAAGAAATATTTTAGGAGTAGGAGGATATGGAATGTGCTATAAAGGATTC	1162
Qy	294	LeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer--GlyGlnGly	312
Db	1163	CTACAAGATGGCACTATAGTAGCAATATAAAGGTTGAAGATGTAATGTATGGGAGGAGGA	1222

Qy	313	Glua	g	l	u	a	r	g	l	u	p	h	e	g	n	a	l	u	a	l	g	l	u	l	e	s	e	r	a	r	g	v	a	l	h	i	s	a	r	g	h	i	s	l	e	u	v	a	l	332												
Db	1223	GAA	A	T	T	C	A	A	T	T	T	C	A	A	C	A	G	A	G	T	G	G	A	A	T	T	T	G	T	G	C	T	G	A	G	A	A	C	C	T	A	T	T	A	1282																	
Qy	333	Ser	L	e	u	V	a	l	G	l	y	T	y	r	C	y	s	t	i	l	e	A	l	a	G	l	y	A	l	a	L	y	s	A	r	g	L	e	u	V	a	l	T	y	r	G	l	u	P	h	e	V	a	l	Pro	352						
Db	1283	CG	A	T	T	G	T	G	A	T	T	T	T	T	G	C	A	C	A	C	T	C	C	A	G	A	G	A	G	G	C	T	T	C	G	T	C	T	A	T	C	C	T	A	C	A	T	G	C	C	A	1342										
Qy	353	Asn	A	s	n	A	s	n	L	e	u	-----	L	e	u	H	i	s	L	e	u	H	i	s	G	l	y	G	l	u	G	l	y	A	r	g	P	r	o	t	h	r	M	e	t	G	l	u	T	r	p	370										
Db	1343	AAT	G	G	A	A	G	T	G	T	G	C	T	T	C	T	T	C	T	T	A	G	A	G	A	T	C	A	T	A	T	A	A	T	G	G	A	A	G	T	T	G	C	C	T	G	C	T	G	A	1402											
Qy	371	Ser	T	h	r	A	r	g	L	e	u	L	y	s	i	l	e	A	l	a	L	e	u	G	l	y	S	e	r	A	l	y	S	e	r	A	l	G	l	y	L	e	u	S	e	r	T	y	r	L	e	u	H	i	s	G	l	u	A	s	p	390
Db	1403	CCT	A	C	T	C	T	C	G	A	A	G	C	T	A	G	C	C	T	T	G	G	A	C	A	G	C	T	A	G	C	T	T	T	G	A	G	G	A	C	T	G	T	A	T	A	T	T	T	T	G	C	A	T	G	A	G	C	A	A	1462	
Qy	391	Cys	A	s	n	P	r	o	b	l	y	S	i	l	e	H	i	s	A	s	p	I	l	e	L	y	s	A	s	e	r	A	s	n	I	l	e	L	e	u	I	l	e	A	s	p	P	h	e	L	y	s	410									
Db	1463	TGT	G	A	T	C	C	A	A	G	A	T	T	A	T	T	C	A	C	G	G	A	T	G	T	G	A	A	G	C	A	A	T	A	T	A	T	T	A	T	T	A	T	T	A	C	T	T	G	G	A	T	G	A	T	1522						
Qy	411	Phe	G	l	u	a	l	a	l	y	s	V	a	l	a	A	s	p	P	h	e	G	l	y	L	e	u	A	l	a	L	y	s	i	l	e	A	l	a	S	e	r	A	s	p	T	h	r	A	s	n	T	h	r	H	i	s	430				
Db	1523	TTT	G	A	A	G	A	C	T	G	T	T	T	G	G	A	T	T	T	T	T	G	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	1582											
Qy	431	Val	S	e	r	T	h	r	A	r	g	V	a	l	M	e	t	G	l	y	T	h	r	P	h	e	G	l	y	T	h	r	L	e	u	A	l	a	P	r	o	G	l	u	T	y	r	A	l	a	A	S	e	r	G	l	y	450				
Db	1583	GT	C	A	C	T	A	C	T	C	T	G	T	T	C	A	G	G	A	C	G	T	A	G	G	C	A	T	G	C	A	T	G	C	C	C	C																									

RESULT 7

US-09-602-472A-3

US-09-002-472A-3
: Sequence 3. Application US/09602472A

; Patent No. 6608240

GENERAL INFORMATION:

APPLICANT: Bidney, Dennis L.

APPLICANT: Hu, Xu

APPLICANT: Li, Guihua

; TITLE OF INVENTION: Sunflower Disease Resistance Genes

FILE REFERENCE: 35718/200630

; CURRENT APPLICATION NUMBER: US/09/602,472A

; CURRENT FILING DATE: 2000-06-23

```
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1559
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)...(1319)
US-09-602-472A-3

Alignment Scores:
Pred. No.: 4,39e-20 Length: 1559
Score: 667.50 Matches: 165
Percent Similarity: 57.50% Conservative: 65
Best Local Similarity: 41.25% Mismatches: 117
Query Match: 19.33% Indels: 57
DB: 4 Gaps: 9

US-10-086-464-2 (1-647) x US-09-602-472A-3 (1-1559)

Qy 219 SerProArgGlnProProProProProProPheMetSerSerSerGlyGly 238
Db 131 TCACCAAAAGTCGGCCATGAAGTCCACCCAT-CCGCCCTT----- 171
Qy 239 SerAspTyrSerAspArgProValLeuProProProSerProGlyLeuVal----- 255
Db 172 -----CCAAATTTCAGATTATCTTCGGGGTGTAGTAGGATGAAGA 213
Qy 255 ----- 255

214 CGAAGAACAAATGTTAATAATGCTGCTAGCTTAAGAGGGAGTCATCAGGCGCGCGGATG 273
Qy 256 LeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPhe 275
Db 274 CTCAA-ATCGCGCGACAAACCTTCACATTCGCGAGCTTGCGCGCCGCCCAATAATTTT 332
Qy 276 SerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeu--- 294
Db 333 CAGCCTGATTGCTTCTTAAGGGAGGGTGTGTTGTTGTTGTTATATAGGAGCGTCTTCAG 392
Qy 295 ProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArg 314
Db 393 AGTAGTGTCAAGTTGTAGCTGTGTTAAACAGCTGTGATAGAAATGCACTCCCAAGGTAACCGT 452
Qy 315 GluPheGlnAlaGluValGluLeuLeuLeuSerArgValHisArgHisLeuValSerLeu 334
Db 453 GAGTTTCTGTTGAAGTTCTTATGCTCAGCCTTTTACATCATCCAAACTTGTGTCATTTG 512
Qy 335 ValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsn 354
Db 513 ATTGCTACTGCTGCGGTGACGAGTACCAAGACTCTTGTCTATGAGTTTATGCGCTTGGGT 572
Qy 355 AsnLeuGluLeuHisLeuHisGly-----GluGlyArgProThrMetGluTrpSerThr 372
Db 573 TCCTTGAAGATCACCTTTCATGATGTTCCGCGCTGACAGAGAACCAATAGATTGG---ACA 629
Qy 373 -ArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAs 392
Db 630 CAGATGAAGATAGCGGCTGT-CCAGCCAGGGGTGGAGTTCTTCATGATAAAGCTAA 688
Qy 392 nProLysIleHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheGly 412
Db 689 CCCACCAAGTTATTCGGGACTCAATCATCAACAATTTGCTGATGAGGATTTCA 748
Qy 412 uAlaLysValAlaAspPheGlyLeuAlaLysIleAla---SerAspThrAsnThrHisVa 431
Db 749 ACCAAAGCTGTCCGACTTCGGGCTCGCGAAACTGGGCCCCACTGGAGCAAGTCTCATGT 808
Qy 431 lSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLy 451
Db ----- 451
```

```
Db 809 TTCACACAGCGGTGATGGGTACATACCGTTACTGTGCTCTCCTGAGTATCCCATGCTGTC 868
Qy 451 sLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuThrGly 471
Db 869 GTTAAACGGTCAAATCCGATGTGTACAGCTTTGGTGTGCTCTTTTACAGCTTATTACCGG 928
Qy 471 YArgArgProValAspAlaAsnAsnValTyrValAspSerLeuValAspTyrAlaAa 491
Db 929 TAGAAAAGCCATAGATAGCACTCAACACATGGACAGACACCTGGTGACATCGGACG 988
Qy 491 qProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMe 511
Db 989 ACCTTTGTTCAAC-----GACAGAAGAAATTCACATCGTTGGTGACCCACCGGTT 1039
Qy 511 tAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysVa 531
Db 1040 AGAAGGTGCGGTACCAATGCGGGGCTGTACAGGCGCTACGCGGTGCATCCATGTGTAT 1099
Qy 531 lArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAs 551
Db 1100 TCAAGAACAGGTTGCAGCCCGGCTTTGATTGCTGTGACGTGTAACGCGCTATCTTATCT 1159
Qy 551 nValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSer-----As 567
Db 1160 TGCA-----AACGAGGGGTAGATCCGAC-CACAGCTCCAGCTTCATATA 1203
Qy 567 nValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAsp 586
Db 1204 CATCATCAGCAGCGCGCGGCAAGGAGAGACTTAAACCGCAAGGTTTTCGAAAAAAC 1261

RESULT 8
US-08-881-706-1
; Sequence 1, Application US/08881706
; Patent No. 6245969
; GENERAL INFORMATION:
; APPLICANT: Chong, Joane
; APPLICANT: Li, Jianming
; TITLE OF INVENTION: Receptor Kinase BIN1
; FILE REFERENCE: 07251/022001
; CURRENT APPLICATION NUMBER: US/08/881,706
; CURRENT FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4104
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (97)..(3687)
US-08-881-706-1

Alignment Scores:
Pred. No.: 4,4e-17 Length: 4104
Score: 604.50 Matches: 161
Percent Similarity: 47.21% Conservative: 84
Best Local Similarity: 31.02% Mismatches: 148
Query Match: 17.51% Indels: 127
DB: 3 Gaps: 14

US-10-086-464-2 (1-647) x US-08-881-706-1 (1-4104)

Qy 107 ProPro-----SerAsnProSerArgGluGlyGlySerProArgProProSer 122
Db 2350 CCACCGGCTAAGTCTTGAACAATCTCGGCTC---TGTTGTTATCTCTTCGCGGTTG 2406
Qy 123 SerProSerProProSerProSerAspGlyLeu-----SerThrGly 137
Db 2407 GATCCTTCA-----AATGCAGACGGTTATGCTCATCATCAGAGATCTCATGGA 2454
Qy 138 ValValValGlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCys 157
Db 2455 AGGAGACCGGCTCCCTTGTGCTGTGAGTGTGCGGATGGATGTGTCTCTTTGTGTGT 2514
```



```
Qy 236 SerGlyGlySerAspTyrSerAspArgProValLeuProProProSerPro-----Gly 253
Db 364 AAGACGGATCGAGGTATCTCTGTAATCTTTGACAGCTCCATCTCTCTCTGCTGGT 423
Qy 254 LeuVal-----LeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAla 269
Db 424 CTTCCTGAGTTTCTCACCTTGGATGGGA---CATTTGGTTCACTCTTAGAGATCTTCAG 480
Qy 270 ArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrVal 289
Db 481 ATGGCTACTAATCATGTTTCAAGGGAATAATCATCGGTGATGCTGGATATGGAGTTGTT 540
Qy 290 HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer 309
Db 541 TACCGCGTAACCTTGTATATGAGTACCTCTGCTGTATAAAGTTGCTCAACAATTTA 600
Qy 310 GlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisArg 329
Db 601 GGACAAGCTGATAAAGACTTCAGAGTTGAAGTTGAAGCTATAGTCACGTTTCGACATAA 660
Qy 330 HisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGlu 349
Db 661 AACTTGGTCCGCTTCTCGGATATTTGATGAAGGAACGCAGAG-----705
Qy 350 PheValProAsnAsnLeuLeuHisLeuHisGlyGlyGluGlyArgProThrMetGlu 369
Db 705 -----705
Qy 370 TrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGlu 389
Db 706 -----CTCGGCTACCTTCAGAG 723
Qy 390 AspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuLeuAspPhe 409
Db 724 CGCATTCAGCCAAAGTGGTGCACAGACATTAAGTCTAGTACATCTCTGATTGATGAC 783
Qy 410 LysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThr 429
Db 784 AAATTCATTTCTAAATTTCTGACTTTGGACTTCTAAACTACTTGGTGTGATGAAGGT 843
Qy 430 HisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSer 449
Db 844 TTTATACTACTAGATTATGGGTACCTTCGTTACGTAGCTCCAGAGTATGCAATTC 903
Qy 450 GlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuLe 469
Db 904 GGTCTTCTGAATGAGAAAGCGATGCTACAGCTTCGGGGTTGACTCTTGGAGCTATA 963
Qy 470 ThrGlyArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrp 489
Db 964 ACTGGTAGATATCCGGTAGACTATGCTCGTCCACCACCCGAGGTACATTTGGTGGAGTGG 1023
Qy 490 AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyGlyLeuAlaAspAla 509
Db 1024 CTGAAGATGATGTC-----CAACAAGACGATCAGAAAGTGGTTGATCCA 1071
Qy 510 LysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaAla 529
Db 1072 AACCTTGAACAAACCAATCTACAAGTGTCTTGAAGAAAGACACTATTGACTGCTTTGAGA 1131
Qy 530 CysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGlu 549
Db 1132 TGTGTTGATCCAATGTCTGAAGAAAGACCGAGGATGAGCCAAAGTTGCACGTATGCTTGA 1191
RESULT 10
US-09-228-986-11
; Sequence 11, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
```

```
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 3097
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-228-986-11
Alignment Scores:
Pred. No.: 4,95e-17 Length: 3097
Score: 600.50 Matches: 173
Percent Similarity: 46.08% Conservative: 80
Best Local Similarity: 31.51% Mismatches: 172
Query Match: 17.39% Indels: 126
DB: 4 Gaps: 13
US-10-086-464-2 (1-647) x US-09-228-986-11 (1-3097)
Qy 66 ProSerThrProGlySerProProPro-----LeuPro 76
Db 1523 CCTGCAAGTTCTGGACGTCCTCAAAACACACGCTTAGACGGAGTTGTTCCCGACGCTTGGG 1582
Qy 77 GlnProSerProAlaProThrThrProGlySerProProAlaProValThrPro--- 95
Db 1583 AGAGCTGAAGACCTTCACCTACTGGAAGTGGAGAAATAACCTGTCTACAAG-GTACCTTAC 1641
Qy 96 -----ProThrArgAsnProProProSerValProGlyProPro 108
Db 1642 CAGATTCCTTGAACACAGACAGCTTGGAGTCCAGAACCTCAGGAACTTGTGCTTTCCT 1701
Qy 109 SerAsnProSerArgGluGlyGlySerProArgProProSerSerProSerPro----- 126
Db 1702 TCTCCACAACCGGTGC-GGTGATGCATCATCTAGTCTTCAATTGAGGACCCCAAGTT 1760
Qy 127 -----ProSerProSerAspGlyLeuSerThrGlyValValValGlyIleAla 143
Db 1761 ACAATAGTTCCCGAGAGAAACAAAGGGGACATAATCGTTTAGCCATTATATCTCGGAC 1820
Qy 144 IleGlyGlyValAlaLeuValIleValThrLeuIleCysLeuLeuCysLysLysLys 163
Db 1821 GTCGGAGGATATCATCTAGTATTTTACTCATCCCGCTCTCGGTATTTCATGTACAGAAG 1880
Qy 164 ArgArgArgAspGluGluAspAlaTyrTyrValProProProProProGlyProLys 183
Db 1881 AGAGGAAGAACTGAA-----1895
Qy 184 AlaGlyGlyProTyrGlyGlyGlnGlnGlnTrpArgGlnAsnAlaThrProPro 203
Db 1895 -----1895
Qy 204 SerAspHisValValThrSerLeuProProProLysAlaProSerProProArgGln 223
Db 1895 -----1895
Qy 224 ProProProProProProProPheMetSerSerGlyGlySerAspTyrSerAsp 243
Db 1896 -----ATGTCATATACGAA 1910
Qy 244 ArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPhe 263
Db 1911 AGGGCAGTCGCAGACGCTGAGAACTCGAAT-----GCAGCTAAGATTTT 1955
Qy 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuGlyGln 283
Db 1956 TCTACAAAGATCAAAACAGCTACAAACACTTTAAAGAA-----GTCATTGGTCAT 2009
Qy 284 GlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLys 303
Db 2010 GGAAGTTTGGATCCGTTGACTTCCAGTGGGNAACCTTCCAGTTGGGNAACCTAGTTGCTGTGAA 2069
```

```
QY 304 GlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeile 323
Db 2070 GTGCGGTTGATAAAACCAACCTTGGTCAGATTCTTTTCATAAATGAGGTTTCGTCTCTTA 2129
QY 324 SerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343
Db 2130 TCACAAGTCGCGCATCAGAACCTTCGTAGCTCGGAAGGATTTTGTATCAGTCGCGAGGT 2189
QY 344 ArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlu 363
Db 2190 CAGATTTTGTCTATGAATATCTACCGGTGATCACTGGCTGATCAACTGTATGTGCA 2249
QY 364 Gly-----ArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAla 381
Db 2250 AACACTAGGAATCTCCTAAGCTGGGTTCGTAGACTCAAGATTGCTGTGATCTGCA 2309
QY 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLys 401
Db 2310 AAAGGACTGGACTATCTACATAATGGAACCAATCTCGAATCATACACGAGACATCAAG 2369
QY 402 AlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAla 421
Db 2370 TGCAGTAATATACTATGGAAGGAGGATGAATGCAAGACTTTGGCGACTTTGGGCTCTCT 2429
QY 422 Lys-----IleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 2430 AAGCAATGATCCAGCCAGACGCA---ACTCAGGTGACCACTGCTGTCAGGGCACAGCT 2486
QY 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 2487 CGTTACCTCGACCTGAATATTAATCTCCACCACTTACAGAGAAAGCGACGCTCTAT 2546
QY 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgGluProValAspAlaAsnAsn 479
Db 2547 ACCTTTGGAGTTGTCTTTTGGAGCTCATCTGTGACGAGAGCCGTTAAATCAITTCAGGA 2606
QY 480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db 2607 ACTCCAGATTCCTCAATTTGGTTTATGGGCAAGCCCTACTTG-----CAG 2654
QY 500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Db 2655 GCAGGTGCAATTTGAG---ATAGTGATGAGATTTTGGGGGAAGTTTCGATGTGGAAGC 2711
QY 520 MetAlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgArgPro 539
Db 2712 ATGAGAAAGTGGCAAAATCCCTGTGAGTCTGTAGAGGGGATGCATCACTAAGGCCA 2771
QY 540 ArgMetSerGlnIleValArgAlaLeuGluGly-----AsnValSerLeuSerAspLeu 557
Db 2772 ACCATTGCACAGATACTGTCTGTCTCAAGAGGCTTACAGCATTCAGCTCTCT----- 2825
QY 558 AsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAsp 577
Db 2826 -----TATCTTGCAGCCTCTGGACATGTGAAC 2852
QY 578 TyrAspSerSerGlnTyrAsnGluAsp 586
Db 2853 TGAATCTACTGTATTACAAACCAAGAC 2879
```

RESULT 11

```
US-08-587-889-1
; Sequence 1, Application US/08587889
; Patent No. 5654397
; GENERAL INFORMATION:
; APPLICANT: CAO, Zhaoan
; APPLICANT: CROSTON, Glenn E.
; APPLICANT: GOEDDEL, David V.
; TITLE OF INVENTION: INTERLEUKIN-1 RECEPTOR-ASSOCIATED
; TITLE OF INVENTION: PROTEIN KINASE AND BINDING ASSAY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
```

```
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,889
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-60916
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3590 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-587-889-1
Alignment Scores:
Pred. No.: 3,48e-16 Length: 3590
Score: 581.50 Matches: 208
Percent Similarity: 38.68% Conservative: 86
Best Local Similarity: 27.37% Mismatches: 239
Query Match: 16.84% Indels: 229
DB: 1 Gaps: 28
US-10-086-464-2 (1-647) x US-08-587-889-1 (1-3590)
```

```
QY 1 MetSerSerAlaProSerPro-GlyThrGlySerProProSerProProSerAsnSerTh 20
Db 80 ATGGCGGGGGCGGGCGGGGAGCCCGCAGCGCCCGGGCCGACAGCTTCTTGAC 139
QY 20 rThrThrThrProProAlaSerAla-----ProProPr 32
Db 140 GAGGTGCGCCCTGGGTTCATGTGCGCTTCTACAAAGTGATGGAGCGCCCTGGAGCCGCGC 199
QY 32 oThrThrProSerProPro----- 39
Db 200 GACTGGTCCGATTGCGCGCCCTGATCGTGCAGCAGACCGAGCTCGGCTGTGCGAG 259
QY 40 ----ProProSerThrIleProThrSerProProSerSerArgSerThrProSerAl 58
Db 260 CGCTCGGGGCGCGCAGCGCGCGCTCTGTGTGCGCTGATCAACCCGAGCCCGGTGTG 319
QY 58 aProProSerProPro-----ThrProSerTh 68
Db 320 GCCGACCTGTGCATCCTCAGCACCTGAGCTGCTCGTGCAGCGGACATCATCACA 379
QY 68 rProGlySerProProLeuProGlnProSerProProAlaProThrThrProGlySe 88
Db 380 GCCTGGCACCTCCCGCCCGCTTCGTCGCCAGCACCACTGCCCGCA-----GG 430
QY 88 rProProAlaProValThrProThrArgAsnProProProSerValProGlyProPr 108
Db 431 CCCAGCAGCATCCTCGACCCCGCG-----AGGCCAGGCTGGAGCCCGCGGA----- 479
QY 108 oSerAsnProSerArgGluGlyGlySerProArgProProSerProSerProSe 128
Db 480 -----AGTTGCCATCTCTCAGCTCCACTTCTC-TCCCCAGCTTT 519
```

Qy 128 rProSerSerAspGlyLeuSerThrGlyValValGlyLeuAlaGlyValAl 148
Db 520 TCCAGGCTCCAGACCCATTCA----- 541
Qy 148 aLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysArgArgAspG1 168
Db 541 ----- 541
Qy 168 uGluAspAlaTyrValProProProProProProGlyProLysAlaGlyProTy 188
Db 542 -----GGGCTGAGCTCGGC----- 556
Qy 188 rGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValva 208
Db 556 ----- 556
Qy 208 lThrSerLeuProProProProLysAlaProSerProProArgGlnProProProPr 228
Db 557 -----CTGGTTCCAGGCCCTGCTTCCCTGCTGGCCTCCAGCCGC 594
Qy 228 oProProProPheMetSerSerSerGlySerAspTyrSerAspArgProValLeuPr 248
Db 595 ATCTCCAGCCCTTCTTCTACCAAGCCAGGC-----CC 627
Qy 248 oProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyGlu----- 266
Db 628 AGAGAGCTCAGTGTCTCTCCAGGAGGCCCGCCCTCTCCGTTTGTGCGCCCTCTG 687
Qy 267 -GluLeuAlaArgAlaThrAnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPh 286
Db 688 TGAGATTTCGGGGCCACCCCAACTTCTCGAGAGCTCAGATCGGGAGGGTGGCTT 747
Qy 286 eGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLy 306
Db 748 TGGGTGCTGTACCGGGCGGTGATG---AGGAACACAGCTGTATCTGTGAAGAGCTGAA 804
Qy 306 aValGlySerGly-----GlnGlyGluArgGluPheGlnAlaGluValGluI1 322
Db 805 GGAGAACGCTGACCTGGAGTGGAGCTGCAGTGAAGCAGAGCTTCTCCAGCAGGTGAGCA 864
Qy 322 eIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAl 342
Db 865 GCTGTCAGGTTTCGTACCCCAACATTTGGACTTGTGCTGGCTACTGTCTCAGAACGG 924
Qy 342 aLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGlu-----LeuHisLe 360
Db 925 CTTCTACTGCTGTGTAGGCTTCTTCTGCCCAACGGCTCCCTGGAGACCGTCTCCACTG 984
Qy 360 uHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuLysE 380
Db 985 CCAGACCCAGCGCTGCCCTCTCTCTGCGCTCAGCGACTGGACATCTCTTCTGGGTAC 1044
Qy 380 rAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp11 400
Db 1045 AGCCCCGGCAATTCAGTTTCTACATCAGGAC---AGCCCCAGCTCATCCATGAGACAT 1101
Qy 400 eLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLe 420
Db 1102 CAAGAGTTCACACGCTCTCTGATGAGAGGCTGACACCCCAAGCTGGGAGACTTTGGCT 1161
Qy 420 uAlaLysIle-----AlaSerAspThrAsnThrHisValSerThrArg-- 434
Db 1162 GGCCCCGTTACCGCGCTTTCAGGCTCCAGGCCCGCAGCAGCAGCATGTGTGGCCGAG 1221
Qy 435 -----ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLe 452
Db 1222 ACAGACAGTGGGGGACCCCTGGCTTCTCCCGAGAGGTACATCAGACGGGAGGCT 1281
Qy 452 uThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyAr 472
Db 1282 GGCTGTGGACACGACACCTTTCAGCTTTGGGGTGTAGTCTAGACACCTTGGCTGTGCA 1341
Qy 472 gArgProVal-----AspAlaAsnValTyrValAspAspSerLeuValAspTr 489

Db 1342 GAGGCTGTGAAGACGACCGTGCCAGGACCAAGTATCTGAAGAC---CTGGTGGAA-- 1396
Qy 489 pAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAl 509
Db 1397 -----GAGGAGGCTGAGGAGGCTGGAGTGGCTTTGAGAGCACCAGAG 1440
Qy 509 aLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAla----- 525
Db 1441 CACACTCAAGCAGGCTGTGGCTGCAGATGCTGGGCTGCTCCCATCGCCATGCAATCTA 1500
Qy 525 ----- 525
Db 1501 CAAGAACCACTGGACCCCGAGCCCGGCGCTGCCACCTGAGCTGGGCTGGGCTGGG 1560
Qy 526 ---CysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnI1 544
Db 1561 CCAGCTGGCTGTGCTGTGCTGCACCGCCGGGCCAAAAGGAGGCTCTCTATGACCCAGT 1620
Qy 544 eValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProG1 564
Db 1621 GTACGAGAGGCTAGAG-----AAGCTGCAGGAGTGTGTGGGGGGTG---CCCGG 1668
Qy 564 Y-----GlnSerAsnValTyrSerSer-- 571
Db 1669 GCATTTGGAGCGCGCCAGCTGCATCTCCCTTCCCGCAGGAGAACTCTTACGTGTCCAG 1728
Qy 572 -----TyrGlyGlySerThrAspTyrAsp----- 579
Db 1729 CACTGGCAGAGCCACAGTGGGGCTGTCTCATGGCAGCCCTGGCAGCGCCATCAGGAGC 1788
Qy 580 -SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnG1 599
Db 1789 CAGTGGCCAGCAGCAGCAGCAGCTGCAGAGAGGCGCCCAACAGCCCGCTGGAGAGTGACGA 1848
Qy 599 uTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSer-- 618
Db 1849 G-----AGCCTAGCGCGCTCTCTGCTGCGCTCGCTCTGGCCTTGACTCCAAGCTG 1902
Qy 619 -----GlySerSerSerGlyGlyGlnThrThrArgG1 629
Db 1903 CCCTCTGGACCCAGCACCCCTCAGGAGGCGCGCTGTCTCAGGGGAGCACGCG-AGGAG 1961
Qy 629 uMetGluMetGly-----LysIleLysArgThrGlyGlnGlyTyrSerGlyPro 645
Db 1962 AATCAGAGCTGGGGAGTGGCGCCAGGATCCCGGCCACAGCCGCTGGAAGGACTGGCCCT 2019

RESULT 12
US-09-016-434-1093
; Sequence 1093, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:


```
Db 557 -----CTGGTTCCAAAGCCCTGCTTCCCTGTGTCCTCCACCGCC 594
Qy 228 oProProPheMetSerSerGlyGlySerAspTyrSerAspArgProValLeuPr 248
Db 595 ATCTCCAGCCCTTCTTACCAAGCCAGGC-----CC 627
Qy 248 oProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGlu----- 266
Db 628 AGAGAGCTCAGTGTCCTCTCAGGAGCGCCCTCTCGTGTTCGTGTCCTCTG 687
Qy 267 -GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPh 286
Db 688 TGAGATTTCCCGGGGACCCCAACATCTTCGAGGAGCTCAAGATCCGGGGGTGGCTT 747
Qy 286 eGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLy 306
Db 748 TGGGTGCGGTATACCGGGCGGTATG---AGGAACACGGGTGATGCTGTGAAGAGGCTGAA 804
Qy 306 sValGlySerGly-----GlnGlyGluArgGluPheGlnAlaGluValGluI 322
Db 805 GGAGAACGCTGACCTGGAGTGACTGCAGTGAAGCAGAGCTTCCTGACCGAGGTGAGCA 864
Qy 322 eIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAl 342
Db 865 GCTGTCCAGGTTTCGTACCAACCAATTTGTGACTTTGTGCTACTGTGCTCAGAACGG 924
Qy 342 aLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGlu-----LeuHisLe 360
Db 925 CTTTACTGCTGTGTGTACGGCTTCCTGCCCAACGGCTCCCTGGAGGACCGTCTCCACTG 984
Qy 360 uHisGlyGlyGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySe 380
Db 985 CCAGACCCAGCGCTGCCACCTCTCTCTGGCTCAGCGACTGGACATCTTCTGGGTAC 1044
Qy 380 rAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIl 400
Db 1045 AGCCCGGGCAATTAGTTTCTACATCAGGAC---AGCCCGACCTCATCTCATGGAGACAT 1101
Qy 400 eLysAlaSerAniLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLe 420
Db 1102 CAAGATTTCCAAAGTCCTTCTGGTGTAGAGGCTGACCCCAAGCTGGGAGACTTTGGCT 1161
Qy 420 uAlaLysIle-----AlaSerAspThrAsnThrHisValSerThrArg--- 434
Db 1162 GSCCGGTTTCAGCGCTTTGCGGGTCCAGCGCCAGCAGCAGCATGGTGGCCCGGAC 1221
Qy 435 -----ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLe 452
Db 1222 ACAGACAGTGGCGGGCACCCTGCGCTTACCTGCCCGAGGAGTACATCAAGACGGGAAGCT 1281
Qy 452 uThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyAr 472
Db 1282 GGCTGTGGACAGGACACCTTCAGCTTTGGGTGGTGTAGTAGACCTTGCTGGTGTCA 1341
Qy 472 gArgProVal-----AspAlaAsnAsnValTyrValAspAspSerLeuValAspTr 489
Db 1342 GAGGCTGTGAAGACCGCAGCGTGCCAGGACCAAGTATCTGAAAGAC---CTGTGTGAA-- 1396
Qy 489 pAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAl 509
Db 1397 -----GAGGAGGCTGAGGAGGCTGGAGTGGCTTTTGAAGACACCCAGAG 1440
Qy 509 aLysMetAsnAsnGlyTyrAspArgGluGluMetAlaAla----- 525
Db 1441 CACTGTCAAGAGGCTGTGGTGCAGATGCCTGGGCTGCTCCCATCGCCATCCAGATCTA 1500
Qy 525 ----- 525
Db 1501 CAAGAAGCACCTGGACCCCGGGCCCTGCCACCTGAGCTGGGCTGGGCTGGG 1560
Qy 526 ----CysAlaAlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGlnIl 544
Db 1561 CCAGCTGGGCTGTGCTGCTGCACCGCGGGGCAAAAGAGGCGCTCTCTATGACCCAGGT 1620
```

```
Qy 544 eValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGl 564
Db 1621 GTACGAGAGGTAGAG-----AAGCTGACGAGCACTGGTGGGGGGTG---CCCGG 1668
Qy 564 Y-----GlnSerAsnValTyrSerSer-- 571
Db 1669 GCATTTGGAGGCGCGCAGCTGCATCCCTTCCCGCAGGAGAACTCTCTAGTGTCCAG 1728
Qy 572 -----TyrGlyGlySerThrAspTyrAsp----- 579
Db 1729 CACTGGCAGAGCCACAGTGGGGCTGCTCCATGGAGCGCCCTTGGCAGGCCCATCAGGAGC 1788
Qy 580 -SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGl 599
Db 1789 CAGTGGCCAGGACAGCAGCAGTGCAGAGAGGCCCAACAGCCCGTGGAGAGTGAAGA 1848
Qy 599 uTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSer-- 618
Db 1849 G-----AGCCTAGCGGCGCTCTCTGCTGCCCTGCGCTCTGCGCACTTTCAGTCCCAAGCTG 1902
Qy 619 -----GlySerSerSerGluGlyGlnThrThrArgGl 629
Db 1903 CCCTCTGGACCCAGCACCCCTCAGGAGCGCGCTGCTCCTCAGGGGGACACCGC-AGGAG 1961
Qy 629 uMetGluMetGly-----LysIleLysArgThrGlyGlnGlyTyrSerGlyPro 645
Db 1962 AATCGAGCTGGGGAGTGGCCCGGATCCCGGCCACACAGCCGTGGAAGACTGGCCCT 2019
```

RESULT 14

```
US-09-602-472A-5
; Sequence 5, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2114
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(2114)
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1475
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1475
; OTHER INFORMATION: n = A,T,C or G
US-09-602-472A-5
```

```
Alignment Scores:
Pred. No.: 2,968-16 Length: 2114
Score: 578.00 Matches: 143
Percent Similarity: 55.07% Conservative: 85
Best Local Similarity: 34.54% Mismatches: 136
Query Match: 16.74% Indels: 51
DB: 4 Gaps: 15
```

US-10-086-464-2 (1-647) x US-09-602-472A-5 (1-2114)

Qy 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProSerProGlyLeu 254
Db 52 GCTGAATCTCTCTCATATTCACCAACCGTGC----- 87
Qy 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274
Db 88 -----CGTCAATTTACCTTTTCGGAGATTCAACTTGCACACCCAAAC 129
Qy 275 PheSerGluAlaAenLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeu 294
Db 130 TTTGATGAGTCGTTGGTGATAGGACGTGGGGGTTGGCAAGGTTTACAGAGGAACCTTC 189
Qy 295 ProSerGlyLys-----GluValAlaValLysGlnLeuLysValGlySerGlyGln 311
Db 190 ACTTATGGGGAACACATCTGGATGCTGCAATTAAGCGCATGGAATCAGGTTCTAGTCAA 249
Qy 312 GlyGluArgGluPheGlnAlaGluValLeuLeuSerArgValHisHisArgHisLeu 331
Db 250 GGACGGTAGAGTTTAGGCTGAATTAAGATGCTCTCAATCTAAGGCACGTGTCATTTG 309
Qy 332 ValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheVal 351
Db 310 GTGCTTTAATTTGTTACTGTAGTATGATGGGCAAGAGATGTTCTGTATATGAACATATG 369
Qy 352 ProHsnAenLeuLeuLeuHisLeuHisGlyGlyArgProThrMetGluTyrSer 371
Db 370 CCCAATGGAACCTTTGCAGATCGCTCCAC---AAGCGTCGAGCTCTCTAACTTGGGTA 426
Qy 372 ThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHis-----Glu 389
Db 427 AGAAGACTCAAAATATGCTAGGAGGGCGCTCGTGGTTAGTACTTGCACACTGGTACG 486
Qy 390 AspCysAenProLysIleIleHisArgAspLysAlaSerAenIleLeuIleAspPhe 409
Db 487 GGTATTAACCATGAGTTATACATCGGAGTGTAAAGAGCAAAATATATTGTTAGATGAC 546
Qy 410 LysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIle-----AlaSerAspThr 427
Db 547 AATGGGCGAGCTAAGGTTTCTGACTTTGGTTTGTCCAAATTTGTCACAAATCAGCCT 606
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db 607 TCACTTATGTTAATCTTTGGTGAAGGACCTTTGGATATATGATGATCGATTTCTT 666
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuGlu 467
Db 667 CAACAGGTAGGCTGACTCGAAAGTCTGACGTGTATGCTTTGGGCGTCTGCTGTTGAA 726
Qy 468 LeuIleThrGlyArgArgProValAlaAsnAenValTyrValAspSerLeuVal 487
Db 727 GTCCATATGTGGGAACAAGTAGTACTGATGAGGAGCAC-TGG-----GGTTTGGCA 776
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 777 ACATGGGCTCAAGACTCTCTT-----AAAGAGGAGGCTTAAGCAAAATGTT 824
Qy 508 AspAlaLysMetAenAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 825 GATTCATAATTTAAGGGGAGAAATATCCCAAAATGTTTGAAGGAGTTTGCACTACTAGCT 884
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db 885 GACCGGTGTTTGTGATAGCCGTCCTCAAGCAACGCTCTCAATGGCGGAGTTGTGATTGGT 944
Qy 548 LeuGluGlyAenValSerLeuSerAspLeuAenGluGly-----Met 561
Db 945 CTTGAGTCAATCTTAGCTTACAGAGAAACCGAGAGTACATGGGTACCAACATTTCTC 1004
Qy 562 ArgProGlyGlnSerAenValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 581
Db 1005 TTTTCCA-----TCCACTTGGGAAAGACAGTTGGAGGTACAGATTTAAATATCCCTG 1055
Qy 582 Gln-----TyrAsn-----GluAsp-----MetLysLysPheArgLysMet 593

Db 1056 GAACGATACCTTACAAATGTTGGAGGTGAAGACAAATAAGTACACAAGTTTGATTTTATA 1115
Qy 594 AlaLeu-----GlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSer 611
Db 1116 ACGATCTTCATGCACTGAACAACTTCTCTGAAGCTAATAAGATTTTCATCTGCTCATAT 1175
Qy 612 AspTyrGlyLeuTyrProSerGlySerSerSerGlyGln 625
Db 1176 GAT---TCCGTGTAC---AAGGGACGGCTACAAAATGGCAA 1211
RESULT 15
US-08-265-628-1
; Sequence 1, Application US/08265628
; Patent No. 5821094
; GENERAL INFORMATION:
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Goring, Daphne
; TITLE OF INVENTION: S-LOCUS RECEPTOR KINASE GENE IN A
; TITLE OF INVENTION: SELF-INCOMPATIBLE BRASSICA NAPUS LINE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 W. Madison St. Suite 3400
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/265,628
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/959,945
; FILING DATE:
; APPLICATION NUMBER: US 07/847,564
; FILING DATE: 03-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pochopien Ph.D., Donald J.
; REGISTRATION NUMBER: 32,167
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-707-8889
; TELEFAX: 312-707-9155
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2749 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Brassica napus
; STRAIN: oleifera
; INDIVIDUAL ISOLATE: W1
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: S-locus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2574
; PUBLICATION INFORMATION:
; AUTHORS: GORING, DAPHNE
; TITLE: THE S-LOCUS RECEPTOR KINASE GENE IN A
; TITLE: SELF-INCOMPATIBLE BRASSICA NAPUS LINE ENCODES A
; TITLE: FUNCTIONAL SERINE/THREONINE KINASE
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2749

US-08-265-628-1

Alignment Scores:

Pred. No.:	5,048-16	Length:	2749
Score:	575.00	Matches:	154
Percent Similarity:	49.80%	Conservative:	94
Best Local Similarity:	30.92%	Mismatches:	160
Query Match:	16.65%	Indels:	91
DB:	1	Gaps:	14

US-10-086-464-2 (1-647) x US-08-265-628-1 (1-2749)

Qy	130	SerSerAspGlyLeuSerThrGlyValValValcglyllealalleGlyGlyValAlaLeu	149
Db	1327	AACCGAATGGGAACAACCATAGCTTGTATTGTTGGAGTTTTGGTT-----	1371
Qy	150	LeuValIleValThrLeuIleCysLeuLeuCysylslyslysArgArGAspGluGlu	169
Db	1372	CTGCTTCATTATGATCATGTTCTGC--CTCTGGAAGAAGAACAAAGCCGACAAACA	1428
Qy	170	AspAlaTyrrTyrrValProProProProGlyProLysAlaGlyGlyProTyrrGly	189
Db	1429	ACTGCAACATCTATTGTAAT-----	1449
Qy	190	GlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValValThr	209
Db	1450	-----CGACAGAGAAAC-----	1479
Qy	210	SerLeuProProProProLysAlaProSerProProArgGlnProProProProPro	229
Db	1480	GGGATG-----ATACTATCAAGCAAGACAGATGGCT-----	1512
Qy	230	ProProMetSerSerSerGlySerAspTyrrSerAspArgProValLeuProPro	249
Db	1513	-----ATAGAGAACAAACTGAGGAATTGAACTTCATTGATA-----	1551
Qy	250	ProSerProGlyLeuValLeuGlyPheSerIysSerThrPheThryrGluGluLeuAla	269
Db	1552	-----GAGTTGGAAGCTGTTGTC-----	1569
Qy	270	ArgAlaThrAnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrrVal	289
Db	1570	AAAGCCACCAGAAAATTTCTCAAATGTAAACAACCTCGACAAAGTGTTTCGGTATTGTT	1629
Qy	290	HisLysGlyValLeuProSerGlyLVysGluValAlaValVallysGlnLeuLysValGlySer	309
Db	1630	TACAAGGGTAGATTACTTGTATGGCAGAGAAATTCGGTAAANAAGGCTATCAAAAACTCG	1689
Qy	310	GlyGlnGlyGluArGluGluPheGlnAlaGluValdluleIleSerArgValHisHisArg	329
Db	1690	GTTCAAGGGACTGGTGAGTTTATGAATGAGGTGAGATTGATCGCAGAGCTTCAGCATATA	1749
Qy	330	HisLeuValSerLeuValGlyTyrrCysIleAlaGlyAlalysArgLeuLeuValTyrrGlu	349
Db	1750	AACCTTGTCGAATTCCTGGCTGTGTGATGAGCAGACAGAGATGCTGGTATATAG	1809
Qy	350	PheValProAnAsnLeuLeuHisLeuHisLysGlyGluGlyArg---ProThrMet	368
Db	1810	TATTTAGAAATTTAAGCTTCGATTCCTATCTTCGAAATAAACGAAGCTCTACGTTA	1869
Qy	369	GluTrpSerThrArgLeuLysIleAlaLeuGlySerAlalysGlyLeuSerTyrrLeuHis	388
Db	1870	AATTGGAAGGACATTTCAACATTTACCAATGGTGTCTCGAGGACTTTTATATCTTCAT	1929
Qy	389	GluAspCysAnProLysIleIleHisArgAspIleLysAlaSerAnilleLeulleAsp	408
Db	1930	CAAGACTCACGGTTTAGGATAATCCACAGAGATATGAAGTAAGTAACATTTTGGCTTGAT	1989
Qy	409	PheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsn	428
Db	1990	AAAAATATGACACCAGAAAGATCTCGGATTTTGGGATGCCGAATCTTTGCAAGGACGAG	2049
Qy	429	ThrHisValSerThrArg---ValMetGlyThrPheGlyTyrrLeuAlaProGluTyrrAla	447

[illegible]

Search completed: April 25, 2004, 01:02:38
Job time : 181 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2004, 00:01:06 ; Search time 638 Seconds
(without alignments)
4572.228 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSSAPSGTSGPPSPNST.....REMEMGKIKRTGQYSGPSL 647

Scoring table:

BLOSUM62
-Q /cgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2907579 seqs, 2254313464 residues

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgapop 10.0 , Xgapext 0.5
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -NATRIX=blom2
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10086464 @CGN 1.1.333 @runat 23042004 083017 5726
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgapop 10.0 , Xgapext 0.5
- 2: /cgapop 10.0 , Xgapext 0.5
- 3: /cgapop 10.0 , Xgapext 0.5
- 4: /cgapop 10.0 , Xgapext 0.5
- 5: /cgapop 10.0 , Xgapext 0.5
- 6: /cgapop 10.0 , Xgapext 0.5
- 7: /cgapop 10.0 , Xgapext 0.5
- 8: /cgapop 10.0 , Xgapext 0.5
- 9: /cgapop 10.0 , Xgapext 0.5
- 10: /cgapop 10.0 , Xgapext 0.5
- 11: /cgapop 10.0 , Xgapext 0.5
- 12: /cgapop 10.0 , Xgapext 0.5
- 13: /cgapop 10.0 , Xgapext 0.5
- 14: /cgapop 10.0 , Xgapext 0.5
- 15: /cgapop 10.0 , Xgapext 0.5
- 16: /cgapop 10.0 , Xgapext 0.5
- 17: /cgapop 10.0 , Xgapext 0.5
- 18: /cgapop 10.0 , Xgapext 0.5
- 19: /cgapop 10.0 , Xgapext 0.5

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
------------	-------	--------------	----	-------------

ALIGNMENTS

RESULT 1

US-10-086-464-1
; Sequence 1, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1944
; TYPE: DNA

Sequence 1, Appli

Sequence 3, Appli

Sequence 95479, A

Sequence 2326, A

Sequence 33030, A

Sequence 33031, A

Sequence 9, Appli

Sequence 9117, Ap

Sequence 13, Appli

Sequence 12, Appli

Sequence 57846, A

Sequence 7339, Ap

Sequence 14127, A

Sequence 6300, Ap

Sequence 47702, A

Sequence 7, Appli

Sequence 6, Appli

Sequence 15, Appli

Sequence 25269, A

Sequence 71008, A

Sequence 35509, A

Sequence 2328, A

Sequence 36420, A

Sequence 42332, A

Sequence 4027, Ap

Sequence 4273, Ap

Sequence 7549, Ap

Sequence 12314, A

Sequence 24475, A

Sequence 30665, A

Sequence 3666, Ap

Sequence 3455, Ap

Sequence 29210, A

Sequence 13131, A

Sequence 1079, Ap

Sequence 1079, Ap

Sequence 311, App

Sequence 25617, A

Sequence 33063, A

Sequence 5, Appli

```
! ORGANISM: Brassica napus
! FEATURE:
! NAME/KEY: CDS
! LOCATION: (1)..(1944)
US-10-086-464-1
Alignment Scores:
Pred. No.: 9,07e-216 Length: 1944
Score: 3453.00 Matches: 647
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-10-086-464-2 (1-647) x US-10-086-464-1 (1-1944)
Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProSerAsnSerThr 20
Db 1 ATGTCCTCGCGCGCTCTCCGGGACTGCTTCGCTCCATCTCCACCATCAAACTCCACA 60
Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProPro 40
Db 61 ACCACCACTCTCTCCAGCTTCCTCCTCTCCACACACACTTCTTCTCTCCGCGG 120
Qy 41 ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaProPro 60
Db 121 CCATCCACTATTCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Qy 61 ProSerProProThrProSerThrProGlySerProProProProLeuProGlnProSerPro 80
Db 181 CCATCTCCACCAACTCCATCTACGCGGGGATCTCCACCTCTCTCTCTCTCTCTCTCT 240
Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100
Db 241 CCCGCTCCAACCTACGCGCGGATCTCCACCGCACCTGTACTCTCTCTCTCTCTCTCT 300
Qy 101 ProProSerValProGlyProProSerAsnProSerArgGluGlyGlySerProArgPro 120
Db 301 CCACCTTCAGTCCAGGACCCGCTCCAATCTCTCCGCGAAGAGGATCTCTCTCGACT 360
Qy 121 ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140
Db 361 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Qy 141 GlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160
Db 421 GGAATCGCATCGGAGGAGTCGCTCTGCTTGATAGTAGACTCTGATTTGTCTCTCTCT 480
Qy 161 LysLysLysArgArgAspGluGluAspAlaTyrTyrValProProProProProPro 180
Db 481 AAGAAGAAACGACGAGAGACGGAAGATGCTTACTATGTTCTCTCGCCACCTCTCTCT 540
Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlnGlnGlnGlnThrArgGlnGlnAsnAla 200
Db 541 GGTCCCAAAGCCGAGGACTTACGCTGACAGCAGCAACAATGGCGGCAACAACGCA 600
Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProProProProPro 220
Db 601 ACACCACTCTCAGATCATGCTGAGCGTCTACCTACCACTCCACCTCAAGGCTCCATCTCA 660
Qy 221 ProArgGlnProProProProProProProPheMetSerSerGlyGlySerAsp 240
Db 661 CCACGGCAACCTCTCCACTCCACCGCTTTTCATGAGCAGCAGCGCGGCTCCGAC 720
Qy 241 TyrSerAspArgProValLeuProProProProProProGlyLeuValLeuGlyPheSerLys 260
Db 721 TACTCGGACCGTCCAGTTCTCTCTCCACCTCTCCAGGGCTTGTGTAGGCTTCTCCAAA 780
Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280
Db 781 AGCACTTTCACATACGAGGAGCTAGTAGAGCCACAATGGTTTCTCCGAGCGCACTTG 840
Qy 281 LeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300
```

```
Db 841 TTAGGACAAAGCGGTTTCGGTTACGTGCACAAAGGTGTGTTCCTAGTGGGAAAGAGTT 900
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Db 901 GCTGTGAAGCAGATTCAAAGTTGGGAGTGGTTCAGGAGAGAGGGAGTTTCAGGACAGAGTT 960
Qy 321 GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
Db 961 GAGATCATCAGCAGAGTTCCACACAGCATCTGGTGTCTCTTGTGGTTATTGTCATCGCC 1020
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360
Db 1021 GGTGCCAAAGATTGCTTGTCTATGAGTTGCTTCTTAACAACAATCTCGAGCTTCACCTC 1080
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db 1081 CATGCGGAGGACGCGCTTACAATGGAATGGAGCACCAAGATTGAAGATTGCTCTTCGATCT 1140
Qy 381 AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 1141 GCTAAAGACCTTCTTATCTTCATGAAGATTGCAATCTCTAAATCATTCACCGTATATC 1200
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 1201 AAGGCTTCAACATATTGATAGATTTCAGATTGAAGCTAAGGTTGCTGATTTCGTCTT 1260
Qy 421 AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 1261 GCTAAGATTGCTTCGATACAAACACGCAATGATATCAACACGTGTGATGGGAACCTTTGG 1320
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 1321 TACTTGGCTCCGGATACGCTGCAAGCGGAAAGCTCACGGAGAGTCTGACGTTTCTCA 1380
Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnVal 480
Db 1381 TTTGGCGTTGTGCTTTTGGAGCTCAATTACTGGACGTCGACCGCTGATGCCAACATGTC 1440
Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500
Db 1441 TATGTAGATGACAGCTTAGTTGACTGGGACGACCAATTCCTTAACCGAGCATCTGAGCAA 1500
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluMet 520
Db 1501 GGAGCTTTTGGAGTTTACGTGATGTCAAAGATGAATATGGGTATGACAGAGGAGATG 1560
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1561 GCTCGCATGTTGCTTGTGCTCGCGCTTGTTCGCCATTTCAGCTCGCGCAGACCTCGC 1620
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1621 ATGAGCCAGATTGTGCGTGGCTTGAAGAAATGATCACTGTCTAGATCTTAAACGAAGG 1680
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 1681 ATGAGACCGAGTCAAGCAATGTATACAGCTCATACGGAGGAGGACCGATATGACTCG 1740
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1741 AGCCAGTCAATGAAGACATGAAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTAC 1800
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1801 AACGCCACGGGTGAGTACAGTAATCCGACCACTGATGAGTCTATGAGCTTACCGCTCTG 1860
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1861 AGCAGCGAGGCGCAACCCACACGCGAAATGGAGATGGGGAAGATTAAAGAGAACCGGTCA 1920
Qy 641 GlyTyrSerGlyProSerLeu 647
```

Db 1921 GGTTATAGTGGACCTTCTCTT 1941

RESULT 2

US-10-086-464-3

Sequence 3, Application US/10086464

Publication No. US20020199218A1

GENERAL INFORMATION:

APPLICANT: GORING, Daphne R. et al.

TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES

FILE REFERENCE: P 25,762-A USA

CURRENT APPLICATION NUMBER: US/10/086,464

CURRENT FILING DATE: 2002-02-28

PRIOR APPLICATION NUMBER: US 10/069,304

PRIOR FILING DATE: 2002-02-19

PRIOR APPLICATION NUMBER: PCT/CA00/00966

PRIOR FILING DATE: 2000-08-18

PRIOR APPLICATION NUMBER: US 60/149,466

PRIOR FILING DATE: 1999-08-19

PRIOR APPLICATION NUMBER: US 60/159,122

PRIOR FILING DATE: 1999-10-13

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 3

LENGTH: 2189

TYPE: DNA

ORGANISM: Brassica napus

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(2189)

OTHER INFORMATION:

US-10-086-464-3

Alignment Scores:

Pred. No.: 1.02e-215 Length: 2189

Score: 3453.00 Matches: 647

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 14 Gaps: 0

US-10-086-464-2 (1-647) x US-10-086-464-3 (1-2189)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20

Db 97 ATGTCCTCGCGCGCTCGCGGACCTGGTTGGCTCCATCCACCATCAAACTCCACA 156

Qy 21 ThrThrThrProProAlaSerAlaProProThrThrThrProSerSerProPro 40

Db 157 ACCACCACTCTCTCCAGCTTCGCTCTCTCTCCACCACTCTCTCTCTCTCTCTCT 216

Qy 41 ProSerThrThrProThrSerProProSerSerSerArgSerThrProSerAlaPro 60

Db 217 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276

Qy 61 ProSerProProThrProSerThrProGlySerProProProProLeuProGlnProSerPro 80

Db 277 CCATCTCCACCACTCCATCTACCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCT 336

Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrThrAsnPro 100

Db 337 CCCGCTCCAACTACCGCGGATCTCCACCGGATCTCTCTCTCTCTCTCTCTCTCT 396

Qy 101 ProProSerValProGlyProProSerAsnProSerArgGluGlyGlySerProArgPro 120

Db 397 CCACCTTCAGTCCCGAGGACCACTCCCAATCTCTCCCGAAGGAGGATCTCTCTCT 456

Qy 121 ProSerSerProProSerProProSerProSerSerSerAspGlyLeuSerThrGlyValVal 140

Db 457 CCATCT 516

Qy 141 GlyThrAlaThrGlyGlyValAlaLeuLeuValThrThrThrThrThrThrThrThr 160

Db 517 GGAATGCCCATCGGAGGAGTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 576

Qy 161 LysLysLysArgArgArgAspGluGluAspAlaTyrTyrValProProProProProPro 180

Db 577 AAGAAGAAACGACGGAGAGAGCAAGAAGATGCTTACTATGTCTCTCGCCACCTCTCTCT 636

Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnTyrArgGlnGlnAsnAla 200

Db 637 GGTCCTCCAAAGCCGAGGACCTTACGGTGGACAGCAACAATGGCGGCAACAAAGCGA 696

Qy 201 ThrProProSerAspHisValThrSerLeuProProProProProProProProPro 220

Db 697 ACACCACTCGTCAATCATGTCTGTGACGTCTACTACCACTACCACTACCACTACCA 756

Qy 221 ProArgGlnProProProProProProProProProPheMetSerSerSerGlyGlySerAsp 240

Db 757 CCACGGCAACCTCTCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCTCCAC 816

Qy 241 TyrSerAspArgProValLeuProProProProProProProProProGlyValLeuGlyPheSerLys 260

Db 817 TACTCGGACCGTCCAGTCTCTCTCCACCTCTCTCCACCTCTCTCCACCTCTCTCCAA 876

Qy 261 SerThrPheThrTyrGluGluLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 280

Db 877 AGCATTTCATACAGAGGAGCTAGCTAGAGCCACCAATGGTTCTCCGAGGCGAATTG 936

Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300

Db 937 TTAGGACAAAGCGGTTTCGGTTACGTGCACAAAGGTGTGTTCCTAGTGGGAAAGATT 996

Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320

Db 997 GCTGTGAAGCAGTTGAAAGTTGGAGTGGTTCAGGAGAGAGGAGTTTCAGGAGAGGTT 1056

Qy 321 GluLeuLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysAlaAla 340

Db 1057 GAGATCATCAGCAGAGTTTCACCAAGCATCTGGTGTCTCTGTGTGTGTGTGTGTGT 1116

Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360

Db 1117 GGTGCCAAAAGATTGCTTGTCTATGATGTTGTCTTAAACAACATCTCGAGCTTCACCT 1176

Qy 361 HisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuLysLeuAlaLeuGlySer 380

Db 1177 CATGGCAGGAGCGGCTTCAATGGAATGGAGCACCAGATTGGAAGATTGCTCTTGGATCT 1236

Qy 381 AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysLeuLeuHisArgAspLe 400

Db 1237 GCTAAGGACTTCTTATCTTCATGAAGATTGCAATCTTAAATCATCTACCGTGATATC 1296

Qy 401 LysAlaSerAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 420

Db 1297 AAGCTTCAAAACATATTGATAGATTTCAAGTTTGAAGCTAAGTTGCTGATTTTGGTCTT 1356

Qy 421 AlaLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 440

Db 1357 GCTAAGATTCTCTCTGATACAAACACCATGATCAACACGTTGATGGGAAACCTTTGG 1416

Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460

Db 1417 TACTTGGCTCGGATATACGCTGCAAGCGGAAGCTCACGGAGAAGTCTGAGGTTTCTCA 1476

Qy 461 PheGlyValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 480

Db 1477 TTTGGCGTTTGTCTTTTGGAGCTCATTTACTGGAGCTCATTTACTGGAGCTGAGCTTGT 1536

Qy 481 TyrValAspAspSerLeuValAspTyrAlaArgProLeuLeuLeuLeuLeuLeuLeu 500

Db 1537 TATGTAGATGACAGCTTAGTTGACTGGGCGACCATTTGCTTAAACCGAGCATCTGAGCAA 1596

Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520

Db 1597 GGAGACTTTGAGGGTTTGTGCTGATGCAAAAGATGATGATGATGATGATGATGATGAT 1656

```
QY 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1657 GCTCGCATGGTGTGCTGCTGGCTTGTGTTCCCAATTCAGCTCGCGCAGACCTCGC 1716
QY 541 MetSerGlnIleValArgAlaLeuGluGlyAenValSerLeuSerAspLeuAsnGluGly 560
Db 1717 ATGAGCCAGATTGTGCTGCTGTAGAGGAATGTATCACTGTCAATCTTAAACGAAGG 1776
QY 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlySerThrAspTyrAspSer 580
Db 1777 ATGAGACCAAGGTCGAAGCAATGTATACAGCTCATACGAGGAAGCACCATTATACATCG 1836
QY 581 SerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1837 AGCCAGTACATGAAGACATGAAGAAGTTTAGGAAATGGCACTTGGAACTCAAGAGTAC 1896
QY 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1897 AACGCCACGGGTGAGTACAGTAAATCCGACCAGTACTATGAGACTGTACCCGCTCTGGTTCA 1956
QY 621 SerSerGluGlyGlnThrArgGluMetGluMetGlyIleIleLysArgThrGlyGln 640
Db 1957 AGCAGCGAGGGCCAAACACACACGCAAAATGGAGATGGGGAAGATTAAAGAAACCGGTGAG 2016
QY 641 GlyTyrSerGlyProSerLeu 647
Db 2017 GGTATAGTGGACCTTCCTT 2037

RESULT 3
US-10-424-599-95479
; Sequence 95479, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95479
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57229C.1
US-10-424-599-95479

Alignment Scores:
Pred. No.: 9e-140 Length: 2451
Score: 2285.50 Matches: 450
Percent Similarity: 80.86% Conservative: 74
Best Local Similarity: 69.44% Mismatches: 97
Query Match: 66.13% Indels: 27
DB: 13 Gaps: 16

US-10-086-464-2 (1-647) x US-10-424-599-95479 (1-2451)
QY 13 ProSerProSerAsnSerThrThrThrThrProProAla-----SerAlaPro 30
Db 246 CCCCTCCCTCCGCT---GCACTCCGTCGTACACTCCGCGGCAACTCCGTCGTCTCT 302
QY 31 ProProThrThrProSerSerProProProProSerThrThrThrProProPro 50
Db 303 CCGCGCGGACACCTTCCTCGCACCTCG-----TCAACTCCCTTCCTCGCCCCCTCG 356
QY 51 SerSerArgSerThrProSerAlaProProProSerProProThrProSerThrProGly 70
Db 357 -----GCGACTCCCTCCGCTTCTCCACCGTCCACTCTCTGCTTCTCTCCACCG 407
```

```
QY 71 SerProProLeuProGlnProSerProProAlaProThrThrThrProGlySerProPro 90
Db 408 TCCACTCCACACCGCGCTCAACTTCGCGGCATCG---ACTTCTCCGCGCTCGCGGCCA 464
QY 91 -----AlaProValThrProProThrArgAsnProProProSerValProGlyProPro 108
Db 465 TCGCACTCGCGCGCTCCGCTCCAGTGGCGCGGAGCGCGAGCAGCACTCCGAGTCCACCG 524
QY 109 SerAsnProSerArgGluGlyGlySerProArgProProProSerSerProSerProProSer 128
Db 525 TCCCGAGCTCGCCCTCTCTCCGTCGGATCGAGACGACCACTCCGTCGCTCCGCTCG 584
QY 129 ProSerSerAspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyGlyValAla 148
Db 595 TCTTCTCTGTCAGTATTTCGACCGTGTGTGGTGGGAATCGCGTGGGGCTGTGGCG 644
QY 149 LeuLeuValIleValThrLeuIleCysLeuLeuCys-----LysLysLysArgArgArgAsp 167
Db 645 GTTCTTCTTGTGTGAGCATTTCTCTGCATATATGTTGCCGGAAGAGAGAGAGAGAGAGTGTAT 704
QY 168 GluGluAspAlaTyrTyrValProProProProProProProProProProProProProPro 186
Db 705 GAAGAG-----TACTATGCTCCGCGCGCGCAACCGCGCGGGGACCTAAA---GATGAT 755
QY 187 ProTyrGlyGlnGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHis 206
Db 756 GCATATGTTGTGTCCTCCCGCCACGTCATGG---CAACACAATGTTCCCTCTCTCAAGATCAT 812
QY 207 ValValThrSerLeuProPro-----ProProLysAlaProSerProProArgGln 223
Db 813 GTGGTCTCAATGATGCTCCAAAGCCATCGGCACCACTGCTCCACCGGCTTATCTCGT 872
QY 224 ProProProProProProPheMetSerSerSerGly-----GlySerAspTyr 241
Db 873 CAACCTCCCGCCACCGCTCTTTCATCAGCAGCAGTGGCGGATCTGGATCAAACTAT 932
QY 242 SerAspArgProValLeuProProSerProGlyLeuValLeuGlyPheSerLysSer 261
Db 933 TCAGSCGTGGAATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 992
QY 262 ThrPheThrThrGluGluLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 281
Db 993 ACATTACGATATGAGGAGTGGCAGCGCAACTGATGCTCTCTCTGATGCCAACCTCTCT 1052
QY 282 GlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAla 301
Db 1053 GGACAAAGGAGATTGGATATGTGCACAGAGAAATCTTCCCAACGCGAGAGAGTGGCA 1112
QY 302 ValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGlu 321
Db 1113 GTGAAGCAATTTGAAGCTGGAAGCGGCGCAAGGGAGCGTGAATTTCAAGCTGAAGTTGAG 1172
QY 322 IleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGly 341
Db 1173 ATAATTAGCGGTGCTCATCAAGCATCTGTTCTTCTTGGTGGATCTGCACTACTGGG 1232
QY 342 AlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHis 361
Db 1233 TCCCAGAGGCTGCTGTTTATGAATTTGTTCCCAACAACACATTTGGAATTCATTTGCAT 1292
QY 362 GlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAla 381
Db 1293 GGAAGAGGCGACCTTACCATTGGTGGCCCAAGACTAAAGAAATTTAGTATCTGCT 1352
QY 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLys 401
Db 1353 AAGGACCTGGCGTATCTTCGATGAGATTGTCTCTTAAGATCATCTCATCTGATATCAAA 1412
QY 402 AlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAla 421
Db 1413 GCTGCAACATCTCTCTGATTTTAAAGTTTGAAGCAAGGTTGAGATTTTGGTCTTGA 1472
QY 422 LysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyr 441
```

```
Db 1473 AAGTTTCTCTGATGTCAATACCATGTTCTACTCGAGTGATGGGACTTTTGGGTAT 1532
Qy 442 LeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPhe 461
Db 1533 TTGGCTCCAGATATGCTTCTAGTGGAAACTGACAGACAATCAGATGTTTCTCTAT 1592
Qy 462 GlyValValLeuLeuGluLeuLeuThrGlyArgArgProValAspAlaAsnAsnValTyr 481
Db 1593 GGAGTCACTGCTCTCGAGTTAATAACCGACGACGCGCAGTCGATATAAAATCAAACTTTC 1652
Qy 482 ValAspAspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGlnGly 501
Db 1653 ATGAGGAGATGTTGGTAGACTGGCTAGGCTTGTCTCACACGAGCTTTGGAAGAGAT 1712
Qy 502 AspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMetAla 521
Db 1713 GATTTGATCTATTATGACCAAGGCTCCAGATGACTATGATCCTAATGAGATGCA 1772
Qy 522 ArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMet 541
Db 1773 CGAATGGTGGCTTCTGCTCGCGCTTGCAATCGTCAATCGGCAAGCGTCGACCAAGGAT 1832
Qy 542 SerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMet 561
Db 1833 AGCCAGGTGTTCTGCGCTCTGGAAGGAGATGCTCTCTAGCAGATCTTAACGAAGGAAT 1892
Qy 562 ArgProGlyGlnSerAsnValTyrSerTyrGlyGlySerThrAspTyrAspSerSer 581
Db 1893 AGACTGGACACAGCATATGATGATCTCAT---GAAAGCTCAGATTATGACACTGCA 1949
Qy 582 GlnTyrAsnGluAspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsn 601
Db 1950 CAGTACAAGGAAGACATCAAAAGTTTCAAAAGATGCAATGGCAATCTCGAGGATATGGT 2009
Qy 602 AlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSer 621
Db 2010 GCAAGCAGTGAAGTACAGTCCGCTACAGTCAAGTCAAGTGTGTTTAAACCCATCAGGCTCAAGT 2069
Qy 622 SerGluGly-----GlnThrArgGluMetGluMetGlyLysIleLysArgThr 638
Db 2070 AGTGAAGCAGACAGCGCCCAACCAAGGGAATGGGAATGGAATGGAAGATGAAG---AAC 2126
Qy 639 GlyGlnGlyTyrSerGlyProSer 646
Db 2127 AATCAAGGTTTCAGTGGGAAGTTCT 2150

RESULT 4
US-10-425-114-2326
; Sequence 2326, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 2326
; LENGTH: 1724
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700209610_FLI
US-10-425-114-2326
```

Alignment Scores:

```
Pred. No.: 6.61e-113 Length: 1724
Score: 1870.00 Matches: 363
Percent Similarity: 79.85% Conservative: 53
Best Local Similarity: 69.67% Mismatches: 87
Query Match: 54.16% Indels: 18
DB: 13 Gaps: 9

US-10-086-464-2 (1-647) x US-10-425-114-2326 (1-1724)
Qy 139 ValValGlyIleAlaIleGlyGlyValAlaLeuValIleValThrLeuIleCysLeu 158
Db 1 GTGCTGGCGGTGGCGGCTCGTGTGCTGTGCTGCTGCGCAGCTTCATCTGCTC 60
Qy 159 LeuCys-----LysLysLysArgArgAspGluAspAlaTyrTyrValProPro 176
Db 61 TGTGTCTCGCAAGAACGCGCGCGCGCGCGCTCAGCAGCTACGATACCG 120
Qy 177 ProProProProGlyProLysAlaGlyGlyProTyrGlyGlnGlnGlnGlnTyrArg 196
Db 121 CCGCCCCCGCG---CCGTACAAGAGGATCCATACGTTGGAAAGTACGATGTTG 174
Qy 197 GlnGlnAsnAla-----ThrProProSerAspHisValValThrSerLeuPro---Pro 213
Db 175 CAGCAAAATGCGGCTCTGCTCCACCCCTGAACATGTGTCAAGATGCACCTTCGCT 234
Qy 214 ProProLysAlaProProProProArgGlnProProProProProProProPheMet 233
Db 235 CCGCAGCATATGCAATCTCTCCACGCGCGCGCGCGCCACACAGCGATGCTA 294
Qy 234 SerSerSerGly-----GlySerAspTyrSerAspArgProValLeuProProSer 251
Db 295 AATAGTAGTGTGATGATCTGTTCTTAATTAATCTGTGTGCGGAGATCTTACCTCCAC 354
Qy 252 ProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAla 271
Db 355 CTTGGTCTGTTCTTGGCTTCTCGAAGAGCACATTACCTACGAAAGTGTGTGAGGGCT 414
Qy 272 ThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLys 291
Db 415 ACTGATGATTCGGAATGATCTCTCTTGGACAGGTGTTTGGTATGTTGCACAGA 474
Qy 292 GlyValLeuProSerGlyLysGluValAlaValLysGlnLysValGlySerGlyGln 311
Db 475 GGATTTGCTGCTAATGGCAAGAGATTTGCTGTAACCAACATGAAAGTGGGAGTGGCGAG 534
Qy 312 GlyGluArgGluPheGlnAlaGluValGluIleSerArgValHisHisArgHisLeu 331
Db 535 GCGAGGCTGATTCAGGCTGAGTTGAGATTATCAGCCGAGTACATCACAACACCTT 594
Qy 332 ValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheVal 351
Db 595 GTGCTTTTGGTATTCATTTCTGGAGGCAAGAGGCTGCTTGTCTATGAGTTTGTGTC 654
Qy 352 ProAsnAsnLeuLeuGluLeuHisHisGlyGluArgProThrMetGluTyrSer 371
Db 655 CCAATAAACATTTGGAAATTCACATTATACGCGCAAGATCGACCAACAATGGAGTGGCCT 714
Qy 372 ThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys 391
Db 715 GCTAGATTAAGATCAGTTTGGGTGCTGCCAAGGTTTAGCTTATCTTCATGAAGACTGC 774
Qy 392 AsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPhe 411
Db 775 CATCCAAGATCATCATCGTGCATAAAGGCATCTAACATTTCTTCTTGCATTTCCAAATTT 834
Qy 412 GluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisVal 431
Db 835 GAAGCTAAGGTTGCTGATTTTGGACTTTGCAAGGTTTCACTACTGATAACAACACCATGTT 894
Qy 432 SerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLys 451
Db 895 TCGACAAGATTAATGGGCACCTTTTGGGTATTTGGCACCTGAGTATGATCTTCTTGGCAAG 954
```

```
QY 452 LeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGly 471
DB 955 CTAACAGAAAATCTGATGATATTTTCCTTCGGAGTCATGCTTCTTGAGCTTATTACTGGG 1014
QY 472 ArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAlaArg 491
DB 1015 CGCGCACCAGTTGACACAAACCCAAACATATATGATGACAGCTTGTTGACTGGGCAAGG 1074
QY 492 ProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMet 511
DB 1075 CCATTACTGATCGGACACTTGAGGATGGTGAATGATGCTTTAGTGATCCTCGGCTG 1134
QY 512 AsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysVal 531
DB 1135 GGAAGGACTTCAATCCTAATCAGATGGCAAGAATGATAGCCTGTGCAGCTGCATGTGTA 1194
QY 532 ArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsn 551
DB 1195 CGCCATTCTGCACGTCGTGCGCCAGCTATGATCAGTCAAGTTGTTGCGGCTTTGGAGGGCAAT 1254
QY 552 ValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSer 571
DB 1255 GGTGCTTTGGAGGACCTTAATCAAGGTGTTGCGCTGGCCATAGCCGCTTCTTTGGTGCA 1314
QY 572 TyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArg 591
DB 1315 TAC---AGCAGCTCCGATTACGATTCGCGCCAGTACACAGGACATGAGAAGTTCAAG 1371
QY 592 LysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSer 611
DB 1372 AAGATGGCATTCACAAACAAC-----TATACAGCAGCAGCAATACAGCGCGCAACCACT 1425
QY 612 AspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluMetGlu 631
DB 1426 GAATATGGACAGATACCGTCTGCATCAAGCAGCGAGGGCCACACAGCAGGAGATGGAG 1485
QY 632 MetGlyLysIleLysArgThrGly-----GlnGlyTyrSerGlyPro 645
DB 1486 TCGGGTGCAATGAGAAGGTGCTACAGTGGTGCTACAGCTCAGGATACAGCGGAGCC 1545
QY 646 Ser 646
DB 1546 TCG 1548
```

RESULT 5

```
US-10-425-114-33030
; Sequence 33030, Application US/10425114
; Publication NO. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33030
; LENGTH: 2880
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17055B11_FLI
US-10-425-114-33030
```

Alignment Scores:
Pred. No.: 2,75e-110 Length: 2880
Score: 1833.00 Matches: 356
Percent Similarity: 79.80% Conservative: 51

```
Best Local Similarity: 69.80% Mismatches: 85
Query Match: 53.08% Indels: 18
DB: 13 Gaps: 9
US-10-086-464-2 (1-647) x US-10-425-114-33030 (1-2880)
QY 150 LeuValIleValThrLeuIleCysLeuLeuCys-----LysLysLysArgArgArgAsp 167
DB 1138 GTGCTGCTCGCAGCTTCATCTGCTCTGCTGTCTCCGCAAGAGCGCGCGCGCGCG 1197
QY 168 GluGluAspAlaTyrTyrValProProProProProGlyProLysAlaGlyGlyPro 187
DB 1198 CGCGCGCTCAGCACTACGGATACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1254
QY 188 TyrGlyGlyGlnGlnGlnGlnArgGlnGlnAsnAla-----ThrProProSerAsp 205
DB 1255 TATGTTGGAACGTACCAGAGTTGG---CAGCAAAATGCGCTCTGCTCCACCCCTGAA 1311
QY 206 HisValValThrSerLeuPro---ProProProLysAlaProSerProProArgGlnPro 224
DB 1312 CATGTGGTCAAGATGCACCTTCGCTCCGCGCAGCATATGCCAATCGTCTCCACAGCG 1371
QY 225 ProProProProProProPheMetSerSerSerGly-----GlySerAspTyrSer 242
DB 1372 CCAGCGCGCGCACCACCGAGCGATGCTAAATAGTAGTGTGATCTGGTTCTTAATTAATCTCT 1431
QY 243 AspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThr 262
DB 1432 GGTGGCGAGATCTTACCTCCACCATCCCTGCTGTGCTGTCTTGGCTTCTCGAAGAGCACA 1491
QY 263 PheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGly 282
DB 1492 TTCACCTACGAAGAGCTGTGTGAGGCTACTGATGATCTCGATGTCTTAATCTCTTGA 1551
QY 283 GlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaVal 302
DB 1552 CAAGTGTGTTTGGCTATGTTTACAGAGGATGTGCTGCCTAATGGCAAGAGATTTCTGTA 1611
QY 303 LysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIle 322
DB 1612 AAACAATTTGAACTGGGAAGTGGCCAGGAGAGCTGAGTTCAGGCTGAGGTTGAGATT 1671
QY 323 IleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAla 342
DB 1672 ATCAGCCGAGTACATCACAAACACCTTGTCTTGTGTGGTGTGCTATGCTATTTCTGAGGC 1731
QY 343 LysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHisGly 362
DB 1732 AAGAGGCTGCTTGTCTATGAGTTTGTCCCAATAACACATTTGGAATTCACATTACATGCG 1791
QY 363 GluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLys 382
DB 1792 AAAGTTCACCCACCAATGGAGTGGCTGCTAGATTAAAGATCAGTTTGGTGCTGCCAAG 1851
QY 383 GlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAla 402
DB 1852 GGTTTAGCTTATCTTCATGAGACTGCCATCCAAAGATCATCCATCGTGACATAAAGGCA 1911
QY 403 SerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLys 422
DB 1912 TCTAAACATTTCTTGTGACTTCCATTTGAAGTAAGTTGCTGACTTTGGACTTTGCAAAAG 1971
QY 423 IleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeu 442
DB 1972 TTCCTACTGATTAACAACACCCATGTTTCGACAGAGTAATGGGCACCTTTGGGTATTGT 2031
QY 443 AlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGly 462
DB 2032 GCACCTGAGTATGATCTTCTGCAAGCTAACAGAAAAATCCGATGTATTTTCTTCCGGA 2091
QY 463 ValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnValTyrVal 482
DB 2092 GTCATGCTCTTTGAGCTTATTACTCGGCGGCGGACGAGTTTGACACAAACCCCAACATATATG 2151
```

```
Qy 483 AspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAsp 502
Db 2152 GATGACAGCTTGGTGTGCTGGCAAGGCCATTACTGATGCGAGCACATTGAGGATGGTGAA 2211
Qy 503 PheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArg 522
Db 2212 TATGATGCTTTAGTGGATCTCGGCTGGGAAGGACTTCAATCTTAATGAGATGGCAAGA 2271
Qy 523 MetValAlaCysAlaAlaLysValArgHisSerAlaArgArgProArgMetSer 542
Db 2272 ATGATACCTGTGCAGCTGATGTAGCCCAATCTGCACGTCGTGCGCCACGATGAGT 2331
Qy 543 GlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArg 562
Db 2332 CAGGTCTGCTGGGCTTTGGAGGGCAATGTGTCTTTGGAGGACCTTAATGAAGGTGTTCGG 2391
Qy 563 ProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGln 582
Db 2392 CCTGGCCATAGCCGCTCTTTGGGTCAATC---AGCAGCTCCGATTACGATTCGGCCAG 2448
Qy 583 TyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAla 602
Db 2449 TACAACGAGGACATGAAGAAGTTCAAGAGATGGCATTCAACACAAC-----TATACC 2502
Qy 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSer 622
Db 2503 AGCAGCCCAATACAGCGCGCCCAACAGTGAATATGACAGATACCGTCTGCATCAAGCAGC 2562
Qy 623 GluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly----- 639
Db 2563 GAGGCGCACAGCAGCAGGAGATGGATCGGTGCAATGAAGAAAGGTGGCTACAGTGGT 2622
Qy 640 -----GlnGlyTyrSerGlyProSer 646
Db 2623 GGCTACAGCTCAGGATACAGCGGAGCCTCG 2652

RESULT 6
US-10-425-114-33031
; Sequence 33031, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33031
; LENGTH: 2881
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17055B12_FLI
US-10-425-114-33031

Alignment Scores:
Pred. No.: 2,75e-110 Length: 2881
Score: 1833.00 Matches: 356
Percent Similarity: 79.80% Conservative: 51
Best Local Similarity: 59.80% Mismatches: 85
Query Match: 53.08% Indels: 18
DB: 13 Gaps: 9

US-10-086-464-2 (1-647) x US-10-425-114-33031 (1-2881)
Qy 150 LeuValIleValThrLeuIleCysLeuLeuCys-----LysLysLysArgArgArgAsp 167
```

```
Db 1139 GTGCTGCTGCCAGCTTCATCTGCCTCTGCTGCTCCGCAAGAGCGCCCGCGCCGCGC 1198
Qy 168 GluGluAspAlaTyrTyrValProProProProGlyProGlyProLysAlaGlyGlyPro 187
Db 1199 CCGCGCGCTCAGCACACTACCGATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1255
Qy 188 TyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 205
Db 1256 TATGGTGGAAAGTACAGAGTTGG---CAGCAAAATCGCGCTCCTGCTCCACCCCTGAA 1312
Qy 206 HisValValThrSerLeuPro---ProProProLysAlaProSerProProArgGlnPro 224
Db 1313 CATGTGCTCAAGATGACACCTTCGCTCCCGCAGCATATGCCAATCTCTCCACAGCGC 1372
Qy 225 ProProProProProProPheMetSerSerSerGly-----GlySerAspTyrSer 242
Db 1373 CCAGCGCGCCACACCGGATGCTAAATAGTAGTGTGGATCTGCTTCTTAATCTCT 1432
Qy 243 AspArgProValLeuProProProProGlyLeuValLeuGlyPheSerLysSerThr 262
Db 1433 GGTGGCGAGATCTTACCTCCACCATCCCTGGTGTCTTCTTGGCTTCTCGAAGACACA 1492
Qy 263 PheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuGly 282
Db 1493 TTCACCTACGAGAGCTTTGAGGCTTACTGATGATCTCGGATCTGATCTCTCTTGA 1552
Qy 283 GlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaVal 302
Db 1553 CAAGTGGTGTGGCTATGTTTCACAGAGGATGCTGCTAATGGCAAGAGATGCTGTA 1612
Qy 303 LysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIle 322
Db 1613 AAACAATGAACTGGGAAGTGGCGGAGAGCGTGAGTTCAGGCTGAGGTTGAGATT 1672
Qy 323 IleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAla 342
Db 1673 ATCAGCGAGTATCATCAACAACACCTTGTGTCTTGGTGGCTATTGCTATTCTGAGAGC 1732
Qy 343 LysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuLeuLeuHisLeuHisGly 362
Db 1733 AAGAGCTGCTGCTGCTATGATTTGTTCCCAATACACATTTGGAATTCACATGCG 1792
Qy 363 GluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLys 382
Db 1793 AAAGTGCACCAACAATGGAGTGGCTGCTAGATTTAAAGATCAAGTTGGGTGCTGCCAAG 1852
Qy 383 GlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAla 402
Db 1853 GGTATTAGCTTATCTTCATGAAGACTGCCATCCAAAGATCATCCATCGTACATAAAGCA 1912
Qy 403 SerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLys 422
Db 1913 TCTAACATCTCTTCTGACTTCCATTTGAAGCTTAAGTGTGCTGACTTGGAGTGCAGAG 1972
Qy 423 IleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeu 442
Db 1973 TTCACTACTGATAACAACACCCATGTTTCACAAGAGTAAATGGGCACCTTTGGGTATTG 2032
Qy 443 AlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGly 462
Db 2033 GCACCTGAGTATGATCTCTTGGCAAGCTAACAGAAAATCCGATGTATTCTTCTCGGA 2092
Qy 463 ValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrVal 482
Db 2093 GTCATGCTTCTGAGCTTATTACTTGGCGCGCGCAGGATGTCACACACCCCAACATATG 2152
Qy 483 AspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAsp 502
Db 2153 GATGACAGCTTGGTGTGCTGCGCAAGGCCATTACTGATGCGAGCACATTGAGGATGGTGA 2212
Qy 503 PheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArg 522
```



```
Db 1092 TTTGGAAATATCATCTTCATGGGAAAAATCTTCCGGTAATGAGGTCTCCCACTAGGTGGC 1151
Qy 375 sileAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysI 395
Db 1152 TATCGCTTAGTCTCGGAAGGACTCGCTTACCTTCACGAAGACTGCCATCTCGGAT 1211
Qy 395 eileHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysVa 415
Db 1212 CATTCACCGCGACATCAAGTCTGCAAAATATTTCTCTGGACTTCAACTTTGATGCTATGGT 1271
Qy 415 lAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgVa 435
Db 1272 GCGTGATTTGGATAGTAGTAAAGTAAACATCTGATAACAACACTCATGTATCTACTCGTGT 1331
Qy 435 lMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLy 455
Db 1332 GATCGGAACCTTCGGATATCTAGCTCCAGATATGCTTCAAGCGGTAAATTAACCGAGAA 1391
Qy 455 sSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgProVa 475
Db 1392 ATCCGATGTTTCTCTTACCGGAGTTATGTTATTTGGAACCTTATAACTGGAACACGACGGT 1451
Qy 475 lAspAlaAsnValTyrValAspAspSerLeuValAspTyrAlaArgProLeuLeuAs 495
Db 1452 TGAT---AATAGCATCACCATGGACGACACCTTAGTAGATTGGGCTCGGCTCTTTATGGC 1508
Qy 495 nAtqAlaSerGluGlnGlyAspPheGlyGlyLeuAlaAspAlaLysMetAsnAenGlyTy 515
Db 1509 TCGCGCGCTAGAAATGGAACCTTTAATGAGCTCGCAGATCGGAGGCTTGAAGGCCAATA 1568
Qy 515 rAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAl 535
Db 1569 CAACCCGCAAGAAATGGCTCGAATGGTGACTTGTGCGGCTGCTAGCATTCGTCATTCGGG 1628
Qy 535 aAtqArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSe 555
Db 1629 GCGTAAACGTCCTCAAGATGACCCAGATAGTAGTAAGCGCTTGAAGGAGAAAGTGTCTTAGA 1688
Qy 555 rAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySe 575
Db 1689 TGCTTTAAACGAAGGTGGAAGCAGGACACACAGTAACTTTACCGGTTCATTTGGAGCAG 1748
Qy 575 rThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLe 595
Db 1749 CTCGGATTATAGTCAGACATCTTACAATTCAGACATCAAGAAATTCAGACAGATAGCTTT 1808
Qy 595 uGlyThrGluGluTyr-----AsnAlaThrGlyGluTyrSerAsnProThrSerAs 612
Db 1809 GTCAGGCCAAGAAATTCACAGTCAGTCACTGTGAAGGAACATCTTAGTAAATGATTTCTAGAGA 1868
Qy 612 pTyrGly 614
Db 1869 TATGGGA 1875
```

RESULT 8

```
US-10-086-464-9
; Sequence 9, Application US/10086464
; Publication No. US20020195218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
```

```
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1939
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-086-464-9
Alignment Scores:
Pred. No.: 5,86e-105 Length: 1939
Score: 1748.50 Matches: 367
Percent Similarity: 69.67% Conservative: 81
Best Local Similarity: 57.08% Mismatches: 141
Query Match: 50.64% Indels: 55
DB: 14 Gaps: 16
US-10-086-464-2 (1-647) x US-10-086-464-9 (1-1939)
Qy 3 SerAlaProSerProGlyThrGlySerProSerProSerProSerAsnSerThrThrThr 22
Db 36 TCTGCTCTCCAAACAAACTCCACCTCTCTCCATCTCCACGCTCTAATACCAATTCAACC 95
Qy 23 ThrProProAlaSerAlaProProThrThrThrProSerSerProProProSer 42
Db 96 ACTCTTCTCCG-----CCGGCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAAGGA 146
Qy 43 ThrIleProThrSerProProSerSerArgSerThrProSerAlaProProProSer 62
Db 147 GACTCATCATCATCGCCACCTCTGATTCCACATCTCCACGCTCCCAAGCTCTCTTAAC 206
Qy 63 ProProThrProSerThrProGlySerProProLeuProGlnProSerProProAla 82
Db 207 CCTCCTAAT---TCCTCTAATAACTCTCTCTCCCTCCGTCACAGGCGGTGGAGGAGAA 263
Qy 83 ProThrThrProGlySerProAlaProValThrProProThrArgAsnProProPro 102
Db 264 AGAGGAATAGGAGAAACAATGGTGCAATGATATCCACGTCACCGGCTCTCTCTCT 323
Qy 103 SerValProGlyProProSerAsnProSerArgGluGlySerProArg-----Pro 120
Db 324 TCT-----CCTCCTCTTAGAGTAATGGAGATAATGGTGTAGCAGATCATCGCA 374
Qy 121 ProSer-----SerProSerProSerPro-----SerSerAsp 132
Db 375 CCAGGAGACTGGAGGCTCTCGCTCAGACACCTCTCTCTAGCGGAGGAGCAGTGA 434
Qy 133 Gly-----LeuSerThrGlyValValGlyIleAlaIleGly 145
Db 435 GGAGGTGGAGGTGGAAGAGTAATACGATACAGCGATCATAGTTGGTGTATTAGTCGA 494
Qy 146 GlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysArg 165
Db 495 GCTGGACTTTTGATGATCGTTCTTATTATTGTGTCTTAGACGCAAAAAGAGAGAAA 554
Qy 166 ArgAspGluAlaAspAlaTyrTyrValProProProProProGlyProLysAlaGly 185
Db 555 -----GACTCCTCTCTAC-----CCTGAACCCCATGAAAGGA 584
Qy 186 GlyProTyrGlyGlnGlnGlnGlnTyrArgGlnGlnAlaThrProProSerAsp 205
Db 585 AACCAATAT-----CAATACTA-TGGAACAACACACACACAACTGCTTCACAGAA 634
Qy 206 HisValValThrSerLeuProProProProLysAlaProSerProPro----- 221
Db 635 -----TTATCCGAATTGGCACCTAAATTCACAAGGCGCAAAACCAACTACTGTGG 688
Qy 222 -----ArgGln-ProProProProProPheMetSerSerSerGly----- 237
Db 689 TTGGGAGGCGGTGGACCATCACCGCTCTCTCTCCGCGATGCTCTACAAGCGGAGAGA 748
Qy 238 -GlySerAspTyrSer-----AspArgProValLeuProProSerProGlyLeuVa 255
Db 749 TTCTTTCATGTACTACTAGGCCCAATCACGCCAGTTTATACCTCTCTCTCTCTCTCTAGC 808
```

```
QY 255 lLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPh 275
Db 809 CCTCGGATTCACAAGAGCATTTTACTTACCAAGAGCTTGGCGTGCACACAGGAGGTT 868
QY 275 eSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysLysGlyValLeuPr 295
Db 869 TACGGATGCTAACCTTTTGGGACAGGAGGATTTGGGTATGTCCATAAAGGAGTCTTGCC 928
QY 295 oSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgL 315
Db 929 TAGCGGGAAGAAGTAGCAGTTTAAGAGTTTAAAGCGGTAGCGACCAAGGAGAGAGGA 988
QY 315 uPheGlnAlaGluValGluIleIleSerArgValHisArgHisLeuValSerLeuVa 335
Db 989 GTTTCACGCTAGGTGCGATATCATTTAGCGGTGTCATCATCGTATCTTGTTCCTTGGT 1048
QY 335 lGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAs 355
Db 1049 TGGATATTGCATAGCTGATGGACAGAGGATTTGGTTATGAGTTTGTTCCTAACAAAC 1108
QY 355 nLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuLy 375
Db 1109 TTTTGAATATCATCTTCATGGGAAATCTTCCGGTAATGGAGTTCTCCACTAGGTTCGG 1168
QY 375 sIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysI 395
Db 1169 TATCGCCTTAGGTGCTCGGAAAGGACTCGTTACCTTCACGAGACTGCCATCTCCGAT 1228
QY 395 eIleHisArgAspIleLysAlaSerAsnIleLeuLeuPheLysPheGluAlaLysVa 415
Db 1229 CATTCACCGCAGACCAAGCTTCGCAATATTTCTTTGGACTTCAACTTTTGATGCTATGT 1288
QY 415 lAlaaspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgVa 435
Db 1289 GCGTGAATTTGATTAGCTTAAGTTAACAATCTGATAACAACACTCATGTATCTACTCGTG 1348
QY 435 lMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLy 455
Db 1349 GATGGGAACCTTCGGATATCTAGCTCCAGAAATATGCTTCAACGGTAAATTAACCGAGNA 1408
QY 455 sSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgProVa 475
Db 1409 ATCCGATGTTTCTCTTACGGAGTTATGTTATTTGAACCTTATACTGGAAAAACGCCGT 1468
QY 475 lAspAlaAsnAnValTyrValAspAspSerLeuValAspTTPAlaArgProLeuLeuAs 495
Db 1469 TGAT---AATAGCATCACCATGGACGACACCTTAGTAGATTGGGCTCGGCCCTCTTATGGC 1525
QY 495 nArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTy 515
Db 1526 TCGCGCGCTAGAAAGATGGAAACTTTAATAGCTCCGAGATCGAGGCTTGAAGGCAACTA 1585
QY 515 rAspArgGluGluMetAlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAl 535
Db 1586 CAACCGCAAGAAATGGCTCGAATGTGACTTGTGCGCTGTAGCATTCGTCATTTCGGG 1645
QY 535 aArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSe 555
Db 1646 GCGTAAACCTCCAAAGATGAGCAGATAGTAGAGCGTTAGAGGAGAAAGTGTCTCTTAGA 1705
QY 555 rAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyLysE 575
Db 1706 TCGCTTAAACGAAGGTGTGAAGCCAGGACACAGTAACGTTTACGGGTCTATTGGGAGCAAG 1765
QY 575 rThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLe 595
Db 1766 CTCGGATTATAGTCAGACATCTTACAATGCAGACATGAAGAAATTCAGACAGATAGCTTT 1825
QY 595 uGlyThrGlnGluTyr-----AsnAlaThrGlyGluTyrSerAsnProThrSerAs 612
Db 1826 GTCGAGCCAAGAAATCCCAGTCAGTGTGTAAGGAAACATCTAGTAATGATTCATGAGA 1885
```

```
QY 612 pTyrGly 614
Db 1886 TATGGGA 1892
RESULT 9
US-10-425-114-9117
; Sequence 9117, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9117
; LENGTH: 1424
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700833950_FLI
US-10-425-114-9117
Alignment Scores:
Pred. No.: 2,26e-104 Length: 1424
Score: 1737.50 Matches: 331
Percent Similarity: 88.73% Conservative: 47
Best Local Similarity: 77.70% Mismatches: 41
Query Match: 50.32% Indels: 7
DB: 13 Gaps: 4
US-10-086-464-2 (1-647) x US-10-425-114-9117 (1-1424)
QY 226 ProProProProProPheMetSerSerSerGly-----GlySerAspTyrSerAsp 243
Db 12 CCCCACCACCCGCTCTCTTTCATCAGCAGCAGTGGCGGATCTGGATCAACAACATTCAGC 71
QY 244 ArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPhe 263
Db 72 GGTGAATTTCTCTCTCTCTCTCTCCAGAAATTTCTATGGGTTCTCTAAGAGCACATTTC 131
QY 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGln 283
Db 132 ACGTATGAGGAGTTGGCAGCGCAACTGATGGCTTCTCTGATGCCAACCTCTCTGGACAA 191
QY 284 GlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLys 303
Db 192 GGAGATTTGGATATGTGCACAGAGAAATCTTCCACGGCAGAGAGTGGCAGTGAAG 251
QY 304 GlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLe 323
Db 252 CAATTTGAAGGCTGGAAGCGGCAAGGGGCGGTGAATTTCCAAAGCTGAAGTTGAGATAAT 311
QY 324 SerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343
Db 312 AGCCGTGTCCATCAACAAGCATCTTGTCTTTGGTTGGATCTATGCATCACTGGGTCGCCAG 371
QY 344 ArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuLeuGluHisLeuHisGlyGlu 363
Db 372 AGGCTGCTGTGTTTATGAATTTGTTCCCAACAACAATTTGGAATTCATTTGGATCGAAAA 431
QY 364 GlyArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGly 383
Db 432 GGGAGACCTACCATGGATTGGCCCAAGACTTAAGAAATTTGCTTTTAGGATCTCTTAAGGA 491
QY 384 LeuSerTyrLeuHisGluAspCysAsnProLysIleLeuHisArgAspIleLysAlaSer 403
Db 403
```


Db	712	GCTGGTCAAGGAGGTGGGAATTGGGGTCCACAGCAACCTGTGTGTGCTCT-----	762
Qy	232	PheMetSerSerGlyGlySerAspTyrSerAspArgProValLeuProProProSer	251
Db	763	-----CACAGTGATGCTTCCAACTTAAACCGGTGAACTGCTATACCGTCACCTCAA	813
Qy	252	ProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAla	271
Db	814	--GCTGCAACTCTTGTGTCAACCAAGACATTTTCATACGATGAACATGTCCATTGCA	870
Qy	272	ThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLys	291
Db	871	ACAGAAAGGTTTCGCTCAGTCAAAATTTCTTAGCAACAGGAGGATTTGGGTATGTTCA	930
Qy	292	GlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGln	311
Db	931	GGAGTCTTGCTAGTGGCAAGAAGTTGCGATGAAGAGTCTTAACTTGAAGTGGACAA	990
Qy	312	GlyGluArgGluPheGlnAlaGluValAlaGluLeuLeuLeuSerArgValHisArgHisLeu	331
Db	991	GGGGAAACGCGAGTTTCAAGCAGAGGTTGATATCATTTAGCCGTGTCATCATCGTCATC	1050
Qy	332	ValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgIleuLeuValTyrGluPheVal	351
Db	1051	GTTTCTCTTGTGGATATTGATCTCTCGTGGTCAAAAGACTTTTGGTATTATGAGTTTATA	1110
Qy	352	ProAsnAsnLeuLeuGluHisLeuHisGlyGluGlyArgProThrMetGluTyrProSer	371
Db	1111	CCTAACCAACACTCTTTGAATTTTCATCTTCATGGAAGGGTCTGCGGTTTGGGATTTGC	1170
Qy	372	ThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys	391
Db	1171	ACAAGAGTGAAGATTGCATTGGGATCAGCTAGAGCCCTTGCATATTTGCGATGAAGACTGT	1230
Qy	392	-----AsnProLysIleIleHisArgAspIle	400
Db	1231	AAGAAAAATCTTTATCTCACATATTTGCATCAGTCACGCTCGCATTTCCACAGAGATATC	1290
Qy	401	LysAlaSerAsnIleLeuIleAspPheGlyPheGluAlaLysValAlaAspPheGlyLeu	420
Db	1291	AAAGCTGCAACATCTCTTTGATTTTTCAGTTTTCAGAACCAAGGTGGCAGATTTTGGATTG	1350
Qy	421	AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly	440
Db	1351	GCTAAGCTATCTCAAGCAACTATCTCATGTCTCCACTCGGCTCATGGGAACTTTTGGAT	1410
Qy	441	TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer	460
Db	1411	TACTTAGCTCCAGAGTATGCATCAAGCGGAAAGTTATCCGACAAATCTGATGTTTCTCA	1470
Qy	461	PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnVal	480
Db	1471	TTTCGGAGTAATGCTCTTTGAGCTCATAAACCGGAAGACCTCTCTCGGATCTAACTGGA	1527
Qy	481	TyrValAspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGln	500
Db	1528	GAAATGGAAGATAGCTTGGTAGATTGGCAAGGCTTTTGTTTGAAGCAGCTCAAGAT	1587
Qy	501	GlyAspPheGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluMet	520
Db	1588	GGAGATTACAACCAATTTGGCTGATCCAGCTCTAGAGCTAAACTACAGTCATCAAGAGATG	1647
Qy	521	AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg	540
Db	1648	GTTCAAAATGGCTTCTGTGACGTGTCAGCAATCAGACATTCAGCAAGAAGACGCGCTAAG	1707
Qy	541	MetSerGln-----IleValArgAlaLeuGlu	549
Db	1708	ATGAGCCAGGTTCAAAAACCTCATACCAGTGTGTTGTTCTATTATTGTACGAGCACTAGAA	1767
Qy	550	GlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyr	569

Db 1768 GGAGATATGCAATGATGATCTTAAGTGGAGGAAACAAGACAGGACAAAGCAGCAGTACTTG 182
 Qy 570 SerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetIysLys 589
 Db 1828 AGCCCCGGGAGCGTGAGCTCAGAGTATGACGCAAGCTCGTACACGGCAGACATGAAAAAG 1887
 Qy 590 PhaArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyGluTyrSerAsnPro 609
 Db 1888 TTCAAGAAACTGGCGTTAGAGAATAAAGAATATCAAAGCAGT---GAATATGGTGGAA-- 1941
 Qy 610 ThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThrArgGlu 629
 Db 1942 ACAAGTGAGTATGCGCTTAAACCCCTTCGTCTCAAGTAGTGAA-----GAA 1986
 Qy 630 MetGluMetGlyLysIleLysArgThrGlyGln 640
 Db 1987 ATGAATAGAGGCTCAATGAACCGCAATCCTCAG 2019
 RESULT 11
 US-10-086-464-12
 ; Sequence 12, Application US/10086464
 ; Publication No. US20020199218A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GORING, Daphne R. et al.
 ; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
 ; FILE REFERENCE: P 25,762-A USA
 ; CURRENT APPLICATION NUMBER: US/10/086,464
 ; CURRENT FILING DATE: 2002-02-28
 ; PRIORITY APPLICATION NUMBER: US 10/069,304
 ; PRIOR FILING DATE: 2002-02-19
 ; PRIORITY APPLICATION NUMBER: PCT/CA00/00966
 ; PRIOR FILING DATE: 2000-08-18
 ; PRIOR APPLICATION NUMBER: US 60/149,466
 ; PRIOR FILING DATE: 1999-08-19
 ; PRIORITY APPLICATION NUMBER: US 60/159,122
 ; PRIOR FILING DATE: 1999-10-13
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 12
 ; LENGTH: 2104
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-086-464-12
 Alignment Scores:
 Pred. No.: 6,31e-101 Length: 2104
 Score: 1687.00 Matches: 368
 Percent Similarity: 63.99% Conservative: 87
 Best Local Similarity: 51.76% Mismatches: 142
 Query Match: 48.86% Indels: 114
 DB: 14 Gaps: 25
 US-10-086-464-2 (1-647) x US-10-086-464-12 (1-2104)
 Qy 1 MetSerSerAlaProSerProGlyThrGlySerProSerProProSerAsnSerThr 20
 Db 52 GTGGATTTCATCTCTGCGCCCTGAAACC---TCAAATGGGACACACCGCTCAACCGGAACA 108
 Qy 21 ThrThrThrProProAlaSerAlaProProThrThrProSerSerProProPro 40
 Db 109 TCG-----CCGTCTAATGAGTCAATCGCGCCCAACACACCGCTTCTTCAACCAACCA 159
 Qy 41 ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaProPro 60
 Db 160 TCATCA-----ATATCTGCTCTCCGCCAGATATCTCCGCTTCTTTTTCACCGCG 210
 Qy 61 ProSerProProThr-----ProSerThrProGlySerProProProLeu 75
 Db 211 CCTGCACCAACGCAAGAAACGTCACCTCCTACATCTCGTCTCTCATCGCGCTGTT 270
 Qy 76 -----ProGln-----ProSerProProAlaPro-----Thr 84
 Db 271 GTAGCTAATCCGTCAACCGAGACTCCAGAGAAATCTTCTCACCTGCACCTGAAGGCTCA 330

US-10-424-599-57846

Alignment Scores:
Pred. No.: 3,3e-91 Length: 2655
Score: 1539.00 Matches: 341
Percent Similarity: 55.19% Conservative: 95
Best Local Similarity: 43.16% Mismatches: 141
Query Match: 44.57% Indels: 213
DB: 13 Gaps: 19

US-10-086-464-2 (1-647) x US-10-424-599-57846 (1-2655)

QY 2 SerSerAlaProSerProGlyThrGlySerProProSerProSerAsnSer----- 19
DB TCATCTCTCCCAACA---AACACTTCCACCCACCATCACCTCCAGTTCCTTCAGCCT 154
QY 20 -----ThrThrThrProProProAlaSerAlaPro---ProProThrThr 34
DB 155 AATCAACACAAACAAACAAACACCAATTCCTGCTCTCTCATCTCTCTGAC 214
QY 35 ProSerSerProPro-----ProProSerThrThrThrSerProProProSerSer 52
DB 215 CTTCTGCTCCACCTCTCTCTCTCCTCCTCAGGAGTGCCATTAAACCCCTCTCTCCATA 274
QY 53 ArgSerThrPro-----SerAlaProProProProProProThrProThrPro 69
DB 275 TCACGTCTCTCTCCACCATCATCTACACCACTTTGCCCCCTCCATCACCACT 334
QY 70 GlySerProProProleuPro-----GlnPro 78
DB 335 GATTCCTCCACCACTATGTCACCTGCATCTCCAAACCCACCACTTACAACTCTCCCCC 394
QY 79 SerProProAlaProThrThrProGlySerProProAlaProValThrProProThrArg 98
DB 395 TCACCAACAGTTACAACTCTCCCGCTCACCACAGCTACCACTCCCTCTCCAGCT 454
QY 99 AnProProProSerValPro-----GlyProProSerAsnProSerArgGluGly 116
DB 455 GAAATCCCACTCTCCCTTCCAAATTTGCTCCACCATCTCCACGACC-----GGT 505
QY 117 SerProArgPro-----ProSerSerPro 124
DB 506 TCCCTCTCTCTCAATCACTCCCAACAATTCACCTCTCTCGACCGAATTTCCGCA 565
QY 125 SerProProSerProSer----- 130
DB 566 TCTCTCTCTCTCCGGCCAAATTTTCAAGGCCACCACTAGAACTCTCCACAGAA 625
QY 130 ----- 130
DB 626 GAAATCCCAACAAACTACTCTTCACATGATCTCTCTCCATCAGTTTCTGAAACTCT 685
QY 130 ----- 130
DB 686 CTAACACTCTCTCTCTGANGTTCCCGCTCCATCCACATGCTTCAACTCTCTCTTCA 745
QY 130 ----- 130
DB 746 GACCTTCAGGATCTTACCTCCAGTCTTTGCTGATCCCCCACTAATAAACAAGTG 805
QY 130 ----- 130
DB 806 GTGGGGGTCCGAAGTGTGCTACCTCTCTCTCAACTGAGAAACCCACTCTAGACCT 865
QY 131 ----- 131
DB 866 ACTAATGATGCTAACAAGTATGTCTTCAACAACAACACACTTCACTTCTGGAGGGTGG 925
QY 135 SerThrGly-----ValValValGlyIleAlaIleGlyValAlaLeuValIle 152
DB 926 AGCACTGGAGGATCTGTGGCTATTTGAATTTGATTTGTTTATTTGCTCTGACCTCTCT 985
QY 153 ValThrLeuIleCysLeuLeuCysLysLysLysArgArgArgAspGluAlaTyr 172

DB 986 GTTATGGCTGTGTGGTTTGCACAGAGAAAGAAAGAGGAAACAGGATCAAGAGGTAGT 1045
QY 173 TyrValProProProProProGlyProGlyProGlyProGlyProGlyGlnGln 192
DB 1046 TATGCTGCTCTCTTCCCA----- 1063
QY 193 GlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuPro 212
DB 1064 -----TTTACCCTCATCCCAATTCAGGTACCTTATTC 1096
QY 213 ProProProLysAlaProSerProProArgGlnProProProProProProProPhe 232
DB 1097 TTGAGCGCCAGTCTCCGGCC-----AACTTT 1123
QY 233 MetSerSerSerGlyGlySerAspTyrSerAspArgProValLeuProProProSerPro 252
DB 1124 TTAGTAGTGGCTCTGGTAGTGAATTT-----GTATATCTCCATCAGAGCT 1171
QY 253 GlyLeuValLeuGlyPheSerSerThrPheThrTyrGluGluLeuAlaArgAlaThr 272
DB 1172 GGTGTGTGA---AGTAGTTCAAGATCATGTTTCAATATGAAGAACTTATTCAAGCTACA 1228
QY 273 AsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGly 292
DB 1229 AATGGGTTTTCAGCACAAAATTTGTTGGAGAGGTGGATTGGCTGTGTATATAAGGT 1288
QY 293 ValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlnGly 312
DB 1289 TTGCTGATAGTAGGAAGAGAGTAGTGTGAACAGCTCAAAATTTGGTGGGCAAGG 1348
QY 313 GluArgGluPheGlnAlaGluValGluIleSerArgValHisHisArgHisLeuVal 332
DB 1349 GAACGCGAATTCAGGCGAGAGTTGAGATTATTAGCCGTGTACATCATCGTCTGGTT 1408
QY 333 SerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValPro 352
DB 1409 TCTTTAGTTGGTTACTGTATATCCGAGCATCAGAGATTGCTGTATATGACTATGTTCC 1468
QY 353 AsnAsnLeuLeuGluLeuHisGlyGluArgProThrMetGlnTrpSerThr 372
DB 1469 AACGATATCTTCAATACCTCTCCCGGTGAATAATACACCGTTCTAGATTGGCTACC 1528
QY 373 ArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsn 392
DB 1529 AGAGTCAAGTTGCTGCTGGTGCAGCTCGTGAATAGTTACTTGCATGAAGACTGTCT 1588
QY 393 ProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysGlu 412
DB 1589 CCACGCAATTTATTCGAGATATTAAAGTCATCAACATCTACTTGTATCTCAACTATGAA 1648
QY 413 AlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSer 432
DB 1649 GCTCAAGTTTCGACTTTGGGCTTCAAAATTTGGCATTTAGATTCAAAATACATGTAAT 1708
QY 433 ThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 452
DB 1709 ACACGTGTAATGGGAACCTTTGGGTACATGGCACCAGAAATATGCGACAAGTGGAAA 1768
QY 453 ThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 472
DB 1769 ACTGAAAAAGTCTGATGTATATTTTGGGGTGTGCTTTTGGAGCTAATTAACAGGTGG 1828
QY 473 ArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAlaArgPro 492
DB 1829 AGCCTGTAGATGCACTCAACCAATTTGGTAGAGCGCTGGTTGAATGGGCTGACCT 1888
QY 493 LeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAspAlaLysMetAsn 512
DB 1899 CTGTTGACGGAAGCACTTGACAACTGAGGACTTTGAAATTTTGGTGGATCCAAAGCTGGG 1948
QY 513 AsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAlaAlaCysValArg 532

Dbb 1949 AAGACTACGATAGAAATGTTTCCGATGATCGAGGCTCCGCGAGCTGTGTACGC 2008
Qy 533 HiserAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552
Db 2009 CACTCATCGGTGAAGAGACACGACGATCAGTCAGGTGTGAGAGCTTTAGATTCTTCGAT 2068
Qy 553 SerLeuSerAspLeuAnGluGlyMetArgProGlyGlnSerAsnValTyrSerTyr 572
Db 2069 GAGTTTACGATCTCAATAACGGAATGAACCGGACAGAGTTCCGGT----- 2116
Qy 573 GlyGlySerThrAspTyrAspSerGlnTyrAsnGluAspMetLysLysPheArgLys 592
Db 2117 -----TTTGATTCCGCGACGACATCTGCACAAATCAAGATGTTAGAGG 2161
Qy 593 MetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAsp 612
Db 2162 ATGGCTTTTGGAGCCCAAGT-----AGTTCGGTTTCTTCAATGAGTCTCAGAGTAGC 2215
Qy 613 Tyr-GlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluMetGluMe 632
Db 2216 TGGAGGAGTA-----GAGATCAGCAGCCACCAACTGTCTTC 2251
Qy 632 tGlyLysIleLysArgThrGlyGlnGly 641
Db 2252 TCCCAAAATATAA---CTGGCATGGGA 2276

RESULT 13

US-10-425-114-7339
; Sequence 7339, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7339
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700652889_FLI
US-10-425-114-7339

Alignment Scores:
Pred. No.: 3,3e-91 Length: 2655
Score: 1539.00 Matches: 341
Percent Similarity: 55.19% Conservative: 95
Best Local Similarity: 43.16% Mismatches: 141
Query Match: 44.57% Indels: 213
DB: 13 Gaps: 19

US-10-086-464-2 (1-647) x US-10-425-114-7339 (1-2655)

Qy 2 SerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSer----- 19
Db 98 TCATCTTCTCAACA---AACACTTCCACCCACCATCAGTCCAGTCTTCTCAGCCT 154
Qy 20 -----ThrThrThrThrProProProAlaSerAlaPro-----ProProThrThr 34
Db 155 AATCAAAACAAACAAACAAACAAACCAATCCCTGCTCTCTCTCTCTCTCTCTGAC 214
Qy 35 ProSerSerProPro-----ProProSerThrIleProThrProThrProProSerSer 52
Db 215 CCTTCTGCTCCACCT 274

Qy 53 ArgSerThrPro-----SerAlaProProProProSerProProThrProSerThrPro 69
Db 275 TCATGTGTCTCTCCACCATCATCTACACCAACCACTTTGCCCCCTCCATCATCAACCT 334
Qy 70 GlySerProProProLeuPro-----GlnPro 78
Db 335 GATTCTCCACCATTCGCACTGCAATCTCCAAACCCACAGTTACAAACATCTCCCCC 394
Qy 79 SerProProAlaProThrThrProGlySerProProAlaProValThrProThrArg 98
Db 395 TCACCAACGATTACAAACATCTCCCTCCACCAACAGTACCAACATCCCTCTCCAGCT 454
Qy 99 AsnProProProSerValPro-----GlyProProSerAsnProSerArgGluGlyGly 116
Db 455 GAAACCTCCACCTTCCCTTCCAAATTTGTCTCCACCATCTCCACAGCC-----GGT 505
Qy 117 SerProArgPro-----ProSerSerPro 124
Db 506 TCCCTCTCTCTCAATCACTCCCAACAAATTCACCTCCCTCTCGACCGATTTCCGCA 565
Qy 125 SerProProSerProSer----- 130
Db 566 TCTCTCTCTCTCCGGCCAAATTTTCCAAAGCCCAACCACTAGAACTCTCTCCAGAGAA 625
Qy 130 ----- 130
Db 626 GAAATATCCACAAAACTACTCTTCCATGCATCTCTCCATCAGTTTCTGAAACTCT 685
Qy 130 ----- 130
Db 686 CTTAAACCTCTCTCTCTGATGTTCCCTCCATCCATGCTTCAACTCTCTCTCTCTCA 745
Qy 130 ----- 130
Db 746 GACCTTCAGGATCTTCACTCCAGCTTCTTGCCTGATCCCCCACTAATAAACAAGTG 805
Qy 130 ----- 130
Db 806 GTGGGGGTCCGAAGGTGCTGCTACCTCTCTTCAACTGAGAAACCCACTGTAGACCT 865
Qy 131 -----SerAspGlyLeu 134
Db 866 ACTAATGATGGTACTAACAGTATGTCTTCAACACACACCTTCCACATTTCTGGAGGTTG 925
Qy 135 SerThrGly-----ValValValGlyIleAlaIleGlyValAlaLeuLeuValle 152
Db 926 AGCACTGGAGGATCTGTGGCTATTGGAATTGTAGTTGTTTATTGCTCTCAGCCTTCT 985
Qy 153 ValThrLeuIleCysLeuLeuCysLysLysLysLysLysLysLysLysLysLysLys 172
Db 986 GTTATGGCTGTGTGTTTGCACAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1045
Qy 173 TyrValProProProProProProGlyProLysAlaGlyGlyProTyrGlyGlyGln 192
Db 1046 TATGCTGCTCTCTCTCA----- 1063
Qy 193 GlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuPro 212
Db 1064 -----TTTACCTCATCCCAATTCAGGTACCTTATTC 1096
Qy 213 ProProProLysAlaProSerProProArgGlnProProProProProProProPhe 232
Db 1097 TTGAGCGCGAGTCTCCGGC-----AACTTT 1123
Qy 233 MetSerSerSerGlyGlySerAspTyrSerAspArgProValLeuProProProSerPro 252
Db 1124 TTAGGTAGTGGCTGTGTAGTATTT-----GTATATTCTCATCAGAGCT 1171
Qy 253 GlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThr 272
Db 1172 GGTGGTGTA---AGTAGTTCAAGATCATGGTTACATATGAGAACTTATTCAAGCTACA 1228
Qy 273 AsnGlyPheSerGluAlaAsnLeuLeuGlyGlyGlnGlyPheGlyTyrValHisLysGly 292

```
Db 1229 AATGGGTTTTTTCAGCACAAAATTTGTTGGAGAGGTGATTTGGCTGTGTTTATAAAGGT 1288
Qy 293 ValLeuProSerGlyLeuValAlaValLysGlnLeuLysValGlySerGlyGlnGly 312
Db 1289 TTGCTGATAGATGGAAGAAGATAGCTGTGAACAGCTCAAAATTTGGTGGGCAAGG 1348
Qy 313 GluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisHisArgHisLeuVal 332
Db 1349 GAACCGCAATTCAGGCGAAGATGAGATTTATAGCCGTGTACATCATCGTCATCTGGT 1408
Qy 333 SerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValPro 352
Db 1409 TCTTTAGTTGTTACTGATATCCGAGCATCAGAGATGCTTGTATATGACTATGTTCTCC 1468
Qy 353 AsnAsnAsnLeuLeuHisLeuHisGlyGluGlyArgProThrMetGluTrpSerThr 372
Db 1469 AACGATATCTTCAATACCATCTCCACGGTGAATAGACAGTCTAGATTGGCCTACC 1528
Qy 373 ArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsn 392
Db 1529 AGATCAGAGTTGCTGTTGGTGCAGCTCGTGAATAGCTTACTTGCATGAAGACTGTGAT 1588
Qy 393 ProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu 412
Db 1589 CCACGCATTTATCATCGAGATATTAAAGTCATCAACATCTCTACTGTATCTCAACTATGAA 1648
Qy 413 AlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSer 432
Db 1649 GCTCAAGTTTTCGGACTTTGGGCTTGCACAAATTTGGCATTAGATTTCAAAATACACATGTAAT 1708
Qy 433 ThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 452
Db 1709 ACACGTGTAAATGGGAACCTTTGGGTACATGGCACCAGCAATATGCGCAAGTGGAAACTT 1768
Qy 453 ThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 472
Db 1769 ACTGAAAAGTCATGATATATTTCTTTGGGGTGTGCTTTTGGAGCTAATTAACAGGTGCG 1828
Qy 473 ArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAlaArgPro 492
Db 1829 AAGCCTGTAGATGCATCTCAACCAATTTGGTATGAGAGCTGGTTGAATGGGCTCGACCT 1889
Qy 493 LeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsn 512
Db 1889 CTGTTTGACGGAAGCACTTGACAATCAGGACITTTGAATTTTGGTGATCCAAGACTGGG 1948
Qy 513 AsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArg 532
Db 1949 AAGAACTACGATAGAAATGAAATGTTTCGGATGATCGAGGCTGCCGAGCCTGTGTACGC 2008
Qy 533 HisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552
Db 2009 CACTCATCGGTGAAGAGACCAGCATAGTACGTAGGTGGTGAGAGCTTTAGATTCCTTGGAT 2068
Qy 553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyr 572
Db 2069 GAGTTTACGGATCTCAATAACGGAATGAACCGGACAGAGTTTCGGTG----- 2116
Qy 573 GlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLys 592
Db 2117 -----TTTGATTCGGCGAGCAATCTGCACAAATCAGAATGTTTATGAGAG 2161
Qy 593 MetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAsp 612
Db 2162 ATGGCTTTTGGAGGCAAGAT-----AGTTCGGTTCCTTCAATGAGTCTCAGAGTAGC 2215
Qy 613 Tyr-GlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluMetGluMe 632
Db 2216 TGGAGGAGTA-----GAGATCACAGCCCAACCACTGCTCTC 2251
Qy 632 cGlyLysLysArgThrGlyGlnGly 641
```

```
Db 2252 TCCCAAAATAAAA---CTGGGCATGGGA 2276
RESULT 14
US-10-425-114-14127
; Sequence 14127, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 14127
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB148-034-DS_FLI
US-10-425-114-14127
```

```
Alignment Scores: 1.09e-84 Length: 2190
Score: 1437.50 Matches: 331
Percent Similarity: 61.90% Conservative: 80
Best Local Similarity: 49.85% Mismatches: 167
Query Match: 41.63% Indels: 87
DB: 13 Gaps: 19
```

```
US-10-086-464-2 (1-647) x US-10-425-114-14127 (1-2190)
Qy 1 MetSerSerAlaProSerProGlyThrGlySerProPro----- 13
Db 2 ATGATGCCCTTCTCCCTCCCTCATCTCGCTCGCCCATCCACCTTCGACGACCCCTT 61
Qy 14 SerProProSerAsnSerThrThrThrThrProProAlaSerAlaProPro----- 31
Db 62 TCTCCTTCTCCTCCTCAATCCTCGACGCGAAAGGAGGTGTGAGCACTA 121
Qy 31 ----- 31
Db 122 AGAGCTATTTAGCAATAATACATACATAGCTATACCTCTTTAGGTCAAGGGTGGAGT 181
Qy 32 ---ProThrThrProSerSerProProProSerThrIlePro----- 45
Db 182 TCTCCTTATTCGGAGGAAGAGCTCCATCCTCATCTACAAGACCAACCAATAAATGGCC 241
Qy 46 ThrSerProProProSerSerArgSerThrProSerAlaProProProSerProProThr 65
Db 242 TCTAGCCACAGTCTCCTCGCCATCATCTCCAAAGAGTCGCCGCCACCAAGAACATCA 301
Qy 66 ProSerThrProGlySerProProProLeuProGlnProSerProAlaProThrThr 85
Db 302 GCTTTGGCGCCAAATCATCTTCA-----AGTCACCTCGCCACCAAGGGTGAATCA 355
Qy 86 ProGlySerProAlaProValThrProProThrArgAsnProProProSerValPro 105
Db 356 --AATTTCATCTCCAGTCCCATCGAAATCGTCTGTCCACCAACCGCCACGACCAAG 412
Qy 106 GlyProProSerAsnProSerArgGluGlyGlySerProArgProProSerSerProSer 125
Db 413 AAAAGTGTAGTACATCTAGCTCTAAAGATGGGAAGCAAGAGTCTGCTTCATCGTCT 472
Qy 126 ProProSerProSerSerAspGlyLeuSerThrGlyValValValGlyIleAlaIleGly 145
Db 473 TCATCGTCCGATATACAGCGCTGTGATCATCCGCGGTGTTGTTGGGTG---GTGGGC 529
```


Qy 146 GlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysLysArgArg 165
Db 530 TTCGCTCTGCTCTTGTCCATCGTGGCGTGGTGTGCTGCGCAAGAAAGAAAGAAACGT 589
Qy 166 ArgAspGluGluAspAlaTyrValProProProProProProGlyPro----- 182
Db 590 -----CCGCCCCCAATGAACATGCCCTTCTACACC 619
Qy 183 ---LysAlaGlyGly-----ProTyrGlyGlyGlnGlnGlnGlnGlnGlnGln 198
Db 620 GACGAGAAAGCAATGTATTATCCCAACGCTGCTGCGCGCTATGTGGCAGCAATAT 679
Qy 199 AsnAlaThrProProSerAspHisValValThrSerLeuProProProProLysAlaPro 218
Db 680 GGCAGCAAC-----GGCAGCATCCCTCCCGCGGG-ATGGCACCA 717
Qy 219 SerProArgGlnProPro-ProProProProProPheMetSerSerSerGlyGly 238
Db 718 TCACGGGGCGGGAACCGCTGTCGCTGCTGCGGCTGCGATGGCGGCGCCCTGAGCGG 777
Qy 238 ySerAspTyrSerAsp-----ArgProValLeuProProProSerProGlyLeuVa 255
Db 778 CGAGATGTAAGTCTCGCGCGCGCGCGCGCTCCCTCCCGCCACCTTCGCGGACGTGGC 837
Qy 255 LeuGlyPheSerLysSerThrPheThrTyrGluLeuAlaAlaArgAlaThrAsnGlyPhe 275
Db 838 GCTGGGGTTCTCCAAAGAGCTCTTCTGCTAGCAGAGCTAGCGGCGAGCGAGCTCGGGCTT 897
Qy 275 eSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuPr 295
Db 898 CTCGTCGCGCACTCTGCGGCGAGCGCGGTTCGGGTACGTGTACAGGCGGTGCTCGC 957
Qy 295 o---SerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluAr 314
Db 958 CGGCAGCGGGAAGAGGTGGCGGTGAAGCAGCTCAAGTCCGGGAGCGGCGAGGGGAGCG 1017
Qy 314 gGluPheGlnAlaGluValGluIleLeuSerArgValHisArgHisLeuValSerLe 334
Db 1018 CGAGTTCCAGCGGAGGTGGAGATCATCAGCGCGGTGCACCGCTCACCTGTGTGCTGCT 1077
Qy 334 uValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAnAs 354
Db 1078 CGTGGGTACTGCAATCGCGGCAACCGAGCGATCTGCTTACGAGTTCGTGGCGCAACA 1137
Qy 354 nAsnLeuGluLeuHisLeuHisGlyGlyArgProThrMetGluTyrSerThrArgLe 374
Db 1138 CACCTGGAGCACCATCTGTACGCCAAGAGCGGCTGTCTATGAGTACGAGCACCAGCAT 1197
Qy 374 uLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLy 394
Db 1198 GAAGATCGCGCTCGGCTCCGCAAGGCGCTCGCTACGAGTTCGTGGCGCAACA 1257
Qy 394 eIleIleHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLy 414
Db 1258 GATCATCCAGCGGACATCAAGCGCGCAACATCTGCTGGCAACAACATTCGAGGGCAT 1317
Qy 414 sValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrAr 434
Db 1318 GGTGGCGGACTTCGGGTGGCCAGCTCAGCAGCGACACCAACAGCAGCTCTCCACGCG 1377
Qy 434 gValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGl 454
Db 1378 CGTCATGGGCACTTCGCTACCTGGCGCGCGAGTACGCGTCCAGCGCAAGCTCACGGA 1437
Qy 454 uLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgPr 474
Db 1438 CAGTCCGAGCGTCTTCTCTTCGCGCATGCTGCTGAGCTACTCAGCGGCGCGCGCC 1497
Qy 474 oValAlaAlaAsnValTyrValAspAspSerLeuValAspTyrAlaArgProLeuLe 494
Db 1498 CATCATACCAACCAAC---TACATGGAGGACAGCTGTGGACTGGGCGCGCGCTGCT 1554
Qy 494 uAsnArg-----AlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAs 512

RESULT 15

US-10-425-114-6300
; Sequence 6300, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 6300
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700560157_FLI
US-10-425-114-6300

Alignment Scores:
Pred. No.: 8,16e-85 Length: 1500
Score: 1437.00 Matches: 284
Percent Similarity: 74.27% Conserved: 45
Best Local Similarity: 64.11% Mismatches: 80
Query Match: 41.62% Indels: 34
DB: 13 Gaps: 7

US-10-086-464-2 (1-647) x US-10-425-114-6300 (1-1500)

Qy 185 GlyGlyProTyrGlyGlyGlnGlnGlnGlnTyrArgGlnGlnAsnAlaThrProProSer 204
Db 11 GGAGTTCACATGGG----- 25
Qy 205 AspHisValValThrSerLeuProProProProLysAlaProSerProProArgGlnPro 224
Db 26 GACCATGTTCTG---AGGATGCAACAAATGGAATGGTCTCTGGTGGCGGAGGTGGGGT 82
Qy 225 ProProProProProProPheMetSerSerSerGly-----GlySerAspTyrSer 242
Db 83 GCACCCACCAACCTCTCTCAATGATGATGAGTAGTCTGAGTTTAGCTCCATCTCATCTCG 142

